

Abstract

Title of Dissertation: MUTATIONAL ANALYSIS OF HUMAN
 IMMUNODEFICIENCY VIRUS TYPE – 1 NUCLEOCAPSID
 PROTEIN TO EVALUATE ITS NUCLEIC ACID CHAPERONE
 ACTIVITY

Nirupama Narayanan, Doctor of Philosophy, 2006

Thesis directed by: Professor Jeffrey J. DeStefano
 Department of Cell Biology and Molecular Genetics

The highly basic 55 amino acid nucleocapsid protein (NC) that coats the HIV-1 genome has two zinc fingers that differ by five amino acids (strain pNL4-3). Previous work showed that NC's first finger (N-terminal) is primarily responsible for unwinding secondary structures (helix destabilizing activity), while the second (C-terminal) plays an accessory role. The amino acid differences between the fingers are (finger one to finger two): phenylalanine to tryptophan (F to W), asparagine to lysine (N to K), isoleucine to glutamine (I to Q), alanine to methionine (A to M), and asparagine to aspartic acid (N to D) at positions 16, 17, 24, 25, and 27 of finger one, respectively. To determine at an amino acid level the reason for the apparent distinction between the fingers, five point mutants were designed with amino acid residues in finger one incrementally replaced by those at corresponding locations in finger two. Each mutant was analyzed in annealing assays with unstructured and

structured substrates. Three groupings emerged: (1) those similar to wild type (wt) levels (N17K, A25M), (2) those with diminished activity (I24Q, N27D), and (3) mutant F16W which had substantially greater helix destabilizing activity than wt NC. All mutants retained wt levels of the condensation/aggregation activity of NC. Unlike I24Q and others, N27D was defective in DNA binding. Only I24Q and N27D showed reduced strand transfer in *in vitro* recombination assays. Double and triple mutants F16W/I24Q, F16W/N27D, and F16W/I24Q/N27D all showed defects in DNA binding, strand transfer, and helix destabilization, suggesting that the I24Q and N27D mutations have a “dominant negative” effect and abolish the positive influence of F16W. Results show that amino acid differences at positions 24 and 27 contribute significantly to finger one's helix destabilizing activity and hence NC's chaperone activity. Preliminary results from *in vivo* experiments indicated that virus with the N27D mutation is infectious at near wt NC levels. This suggests that aggregation activity may be more important than helix destabilizing for viral viability. Results from two other forms of HIV-1 NC (NC_{p9} and NC_{p15}) and NC proteins from Simian Immunodeficiency Virus and Murine Leukemia Virus are also reported.

MUTATIONAL ANALYSIS OF HUMAN IMMUNODEFIENCY VIRUS TYPE – 1
NUCLEOCAPSID PROTEIN TO EVALUATE ITS NUCLEIC ACID CHAPERONE
ACTIVITY

By

Nirupama Narayanan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:

Dr. Jeffrey J. DeStefano, Chair
Dr. William J. Higgins, Dean's Representative
Dr. Richard C. Stewart
Dr. Siba K. Samal
Dr. Vikram N. Vakharia

©Copyright by
Nirupama Narayanan
2006

Dedication

This thesis is dedicated to my loving parents (Narayanan and Akila) and to my dear husband
Sri for all their love and encouragement.

Acknowledgements

My journey through graduate school would not have been possible without significant support and encouragement from several people. I would first like to express my utmost gratitude to my advisor Dr. Jeffrey DeStefano. It was his excellent guidance and constant encouragement that got me through this arduous journey. One thing that I admire about him amongst his many magnificent qualities is his tremendous perseverance. He never gives up easily. Whenever I was about to give up on something, he would convince me to try a few other things, one of which would invariably end up working. I hope to imbibe this quality from him. He is a great person to work with, friendly, kind, extremely patient, and always approachable. He always had the time to answer the silliest of my questions. Thanks a ton, Jeff!

Next, I would like to thank the members of my dissertation committee. I would like to thank Dr. Rick Stewart for giving very valuable suggestions about my fluorescence assays and kinetics data. He instilled in me the importance of getting my basics strong. Dr. Samal constantly encouraged me to develop good knowledge about virology in general, and not just restrict it to retroviruses. My first laboratory rotation was with Dr. Vik Vakharia. The short cloning project that I completed under him proved to be very useful during my thesis work later on. I would finally like to thank Dr. Higgins for making me realize the importance of good writing and oratorical skills.

I am indebted to Dr. Beth Gantt and other members of the recruiting committee in the MOCB program for giving me the opportunity to pursue a PhD, a dream that I had always cherished! Thanks a lot Nancy and Lori for helping me with many administrative details.

I had some really nice people to work with during my PhD. I am extremely grateful to Megan who helped me a lot when I first came into the lab. I know I must have pestered you a zillion times with my silly questions, but you had a patient ear for all of them. I would like to thank Suchi for all the help that she gave me during my first couple of years in the lab. You used to inspire me a lot Suchi: I admire the simple way that you had about yourself and your selfless attitude towards helping others. I especially enjoyed all the coffee breaks with you at the Union! I thank Bill for being such a good lab technician and for helping me with a lot of computer related queries. I thank Aminata and Nusrat for keeping all the glassware clean. I would like to thank my fellow graduate students who helped me a lot during my initial years and with whom I had a lot of fun: Danielle, Marie, Shruti, Athi, Arun, Kavita, Chris, Guo Hong and several others. Many thanks to Shock, Indra, Guru, Arunchandar, Om and Sada for helping us with moving and accommodation. Thanks to the “DESI” student organization in UMD through which I got to meet Sri. Thanks a lot Rajiv and Tejas for helping me during my first days in Maryland.

Finally, I would like to thank my family for all their love and support. I wish to express my utmost gratitude towards my dear parents (Amma and Appa) for their unconditional love and support. They gave me nothing less than the best, believed in me, and encouraged me. Their sincere prayers, blessings, and endless sacrifices have enabled me to get through difficulties in life. I would like to express my gratitude towards my dear brother, Ranga (Babu). It was due to him that I got the opportunity to experience graduate school in the US. He helped me a great deal with the nuances of the application procedure. He helped me choose the right schools to apply to, and the good programs in those schools. I am constantly reminded of my humorous sister-in-law, Lathu whose jokes lightened me up when I needed

to be cheered up. I am indeed very grateful to my sister, Neeraja and my brother in law, Rajesh for their help when I first came here. They received me at the airport and made sure that I had a very relaxing summer break before my classes started. They ensured that I had everything that I might need to start living comfortably on my own. I had lots and lots of fun playing with my darling niece Kavya and naughty little nephew Adi. I always had the comfort that I had family around in case I needed any help. Thanks Babu and Neeraj for watching out for me always!

Next, I wish to express my sincere gratitude towards the one person without whom I would not be graduating today; my dear husband Sri. You never doubted me once. Without your constant encouragement and support, I would have never had the courage to go on and would have slided out a long time back. You are such a wonderful human being! You are the epitome of patience. You always sit down with me and listen to all the things that I complain about.

Finally, and most importantly, I would like to thank the Almighty for blessing me with all the wonderful things in life that I could ever imagine and more.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
LIST OF ABBREVIATIONS.....	x
Chapter 1 General Introduction.....	1
1.1 Introduction.....	1
1.2 AIDS prevalence and general epidemiology.....	2
1.3 Pathogenesis of HIV infection.....	4
1.4 Anti-retroviral therapy.....	5
1.5 HIV morphology and genome organization.....	8
1.6 Viral life cycle.....	13
1.7 Reverse transcriptase.....	18
1.8 Recombination in HIV.....	22
1.9 Nucleocapsid protein.....	26
Chapter 2 Evaluation of the chaperone activity of HIV-1 nucleocapsid protein at an amino acid level.....	34
2.1 Introduction.....	34
2.2 Materials.....	43
2.3 Methods.....	45
2.4 Results of Fluorescence resonance energy transfer (FRET) and gel shift assays with wt NC and NC mutants.....	53
2.5 Results of binding affinity analysis for binding of NC and NC mutants to nucleic acid substrates.....	83
2.6 Results of internal strand transfer assays with wt NC and NC mutants.....	91
2.7 Discussion.....	106
Chapter 3 Differential binding of amino acid residues of HIV-1 NC to nucleic acid templates.....	112
3.1 Introduction.....	112
3.2 Materials.....	116
3.3 Methods.....	117
3.4 Results.....	118
3.5 Discussion.....	127
Chapter 4 General Discussion.....	130
BIBLIOGRAPHY.....	135

LIST OF TABLES

Table 1-1: The nine genes of HIV-1.....	12
Table 2-1: Rate constant (k) calculation for NC proteins on 0.0dna substrate.....	63
Table 2-2: Rate constant (k) calculation for NC proteins on 5.8dna substrate.....	64
Table 2-3: Rate constant (k) calculation using varied [NC] on 5.8dna substrate.....	65
Table 2-4: Rate constant calculation for NC proteins on 9.0dna substrate.....	66
Table 2-5: Sequence of forward and reverse primers used for PCR amplification of the gag-pol region.....	105
Table 3-1: Amino acid sequences of various NC proteins used in this study.....	123
Table 3-2: Rate constant calculation for NC proteins on 0.0dna substrate.....	124
Table 3-3: Rate constant calculation for NC proteins on 5.8dna substrate.....	125
Table 3-4: Rate constant calculation for NC proteins on 9.0dna substrate.....	126

LIST OF FIGURES

Figure 1-1: Schematic of mature HIV virion.....	10
Figure 1-2: HIV-1 RNA genome and proviral DNA.....	11
Figure 1-3: Schematic of Reverse transcription.....	16
Figure 1-4: Viral life cycle.....	17
Figure 1-5: HIV-1 RT.....	20
Figure 1-6: Linear diagram of HIV-1 RT.....	21
Figure 1-7: Schematic of HIV-1 pNL4-3 nucleocapsid protein.....	29
Figure 1-8: HIV-1 NC-SL3 ψ RNA complex.....	31
Figure 2-1: Schematic diagram of FRET.....	55
Figure 2-2: Schematic model of FRET assay.....	56
Figure 2-3: 42mer DNA substrates used for annealing assays.....	58
Figure 2-4: FRET assay with NC finger mutants on 0.0dna.....	61
Figure 2-5: FRET assay with NC finger mutants on 9.0dna.....	62
Figure 2-6: Effect of NC point mutants on annealing of 0.0dna.....	71
Figure 2-7: Effect of NC point mutants on annealing of 5.8dna.....	72
Figure 2-8: FRET assay with NC point mutants on 9.0dna.....	73
Figure 2-9: FRET assay of F16W NC mutant on 9.0dna.....	74
Figure 2-10: Effect of NC point mutants on annealing of 15.8dna.....	75
Figure 2-11: Semi-log plot of the intensity profile of N27D mutant.....	76
Figure 2-12: Effect of NC double and triple mutants on annealing of 0.0dna.....	77
Figure 2-13: Effect of NC double and triple mutants on annealing of 5.8dna.....	78
Figure 2-14: Effect of NC double and triple mutants on annealing of 9.0dna.....	79
Figure 2-15: Autoradiograms of annealing assays on 7.5rna with NC mutants.....	81
Figure 2-16: Autoradiograms of annealing assays done with 9.0dna.....	82
Figure 2-17: Schematic diagram of nitrocellulose filter binding assay.....	85
Figure 2-18: Binding of SSHSd to 0.0dna detected by filter binding assay.....	86
Figure 2-19: Binding of F16W, N17K, I24Q, and A25M NC mutants to 0.0dna.....	87
Figure 2-20: Binding of N27D, F16W/I24Q, F16W/N27D, and F16W/I24Q/N27D NC mutants to 0.0dna.....	88
Figure 2-21: Binding of N27D to 16.3dna.....	89
Figure 2-22: Binding of I24Q to 16.3dna.....	90
Figure 2-23: Schematic model of strand transfer assay.....	94
Figure 2-24: Schematic of experimental approaches used.....	95
Figure 2-25: Predicted secondary structure of <i>gag-pol</i> donor RNA template.....	96
Figure 2-26: Predicted secondary structure of <i>gag-pol</i> acceptor RNA template.....	97
Figure 2-27: Autoradiogram of strand transfer assay with N17K and A25M.....	98
Figure 2-28: Autoradiogram of strand transfer assay with N27D and I24Q.....	99
Figure 2-29: Autoradiogram of strand transfer assay with F16W/I24Q.....	100
Figure 2-30: Autoradiogram of strand transfer assay with F16W/I24Q/N27D and F16W/N27D.....	101
Figure 2-31: Graph showing efficiency of strand transfer <i>versus</i> time for F16W/I24Q, F16W/I24Q/N27D and F16W/N27D.....	102
Figure 2-32: Graph showing efficiency of strand transfer <i>versus</i> time for N17K, A25M and F16W.....	103

Figure 2-33: Graph showing efficiency of strand transfer <i>versus</i> time for N27D and I24Q.....	104
Figure 3-1: FRET assay with retroviral NC proteins on 0.0dna.....	120
Figure 3-2: FRET assay with retroviral NC proteins on 5.8dna.....	121
Figure 3-3: FRET assay with retroviral NC proteins on 9.0dna.....	122

LIST OF ABBREVIATIONS

AIDS	Acquired Immunodeficiency Syndrome
HIV	Human Immunodeficiency Virus
CDC	Centers for Disease Control and Prevention
HTLV	Human T-Lymphotropic Virus
LAV	Lymphadenopathy Associated Virus
ICTV	International Committee on Taxonomy of Viruses
CRF	Circulating recombinant form
OI	Opportunistic infections
PCP	Pneumocystis carinii pneumonia
RT	Reverse transcriptase
RNaseH	Ribonuclease H
<i>gag</i>	group specific <i>antigen</i>
MuLV	Murine Leukemia Virus
DLS	Dimer Linkage Site
PBS	Primer Binding Site
ORF	Open Reading Frame
PPT	Polypurine tract
GDP	Guanosine diphosphate
LTR	Long Terminal Repeats
IN	Integrase
CA	Capsid
PR	Protease

NC	Nucleocapsid
TAR	<i>trans</i> -activation response
FRET	Fluorescence Resonance Energy Transfer
dNTPs	Deoxynucleotide triphosphates
rNTPs	Ribonucleoside triphosphates
AMP	Adenosine mononucleoside phosphate
CIP	Calf intestinal alkaline phosphatase
EDTA	Ethylenediaminetetraacetic acid
PCR	Polymerase chain reaction
wt	wild-type
ss	single-stranded
ds	double-stranded
nt	nucleotide
FAM	fluorescein-6-carboxamidohexyl
DABCYL	4-[[[(4-dimethylamino) phenyl]-azo] benzenesulfonicamino
DTT	Dithiothreitol
PNK	Polynucleotide kinase

Chapter 1 General Introduction

1.1 Introduction

Acquired Immunodeficiency Syndrome or AIDS is one the leading causes for worldwide mortality since it was first reported in 1981. AIDS is caused by the Human Immunodeficiency Virus or HIV which is a member of the retrovirus family (*Retroviridae*). Retroviruses contain a highly divergent and large population of enveloped, positive sense RNA viruses with several common features including viral replication, structure and composition. They are so named because of a characteristic feature in their replication: retroviruses make a double stranded DNA copy of their single stranded RNA genome, contrary to the classical flow of genetic information from DNA to RNA. HIV belongs to the *Lentivirus* genus, one of seven genera in the family. Lentiviruses are otherwise called slow viruses primarily due to their slow growth and long asymptomatic period between initial infection and appearance of detectable clinical symptoms [1]. A person is said to be suffering from AIDS when he/she is in the most advanced stages of HIV infection. The Centers for Disease Control (CDC) defines AIDS as a condition when the T4 lymphocyte count drops below 200 per cubic millimeter of blood. CDC's definition also encompasses to some 26 opportunistic infections which are otherwise harmless to healthy individuals [2]. During May 1983 Luc Montagnier and his co-workers at the Pasteur Institute in France reported that they had isolated a new retrovirus associated with AIDS which they called Lymphadenopathy Associated Virus (LAV). In April of the following year Dr. Robert Gallo of NCI reported that his group had determined the causative agent of AIDS and named it Human T-Lymphotropic Virus Type III (HTLV-III). Subsequently in June of that year, the two groups announced that LAV and HTLV-III were almost identical. Two years later, the International

Committee on Taxonomy of Viruses (ICTV) changed the name of the AIDS virus to Human Immunodeficiency Virus (HIV) [3]. There are two main types of the virus, HIV-1 and HIV-2. HIV-2 was first isolated in Cameroon in West Africa in 1985. It is less prevalent and virulent than HIV-1, though still fatal. The genome of HIV-2 is very similar to SIV or Simian Immunodeficiency Virus [4]. HIV-1 is by far the more predominant form worldwide. There are three main groups of HIV-1: Group M or “Main” is highly predominant and is responsible for more than 90% of infections worldwide. It is comprised of nine different subtypes: A, B, C, D, F, G, H, J, K and various circulating recombinant forms (CRFs) that were derived from recombination between subtypes. Previously, subtypes E and I were also categorized under the M Group. The subtypes E and I are no longer designated as two pure subtypes. They are now being referred to as CRFs. Genetic hybridization between subtypes A and a “parent subtype E” is thought to have resulted in CRF A/E. A pure isolate of subtype E has never been reported. Subtype I was initially so named when it was first isolated in Cyprus. It was then reclassified as a CRF namely CRF A/G/I. It is now believed that this virus represents an even more complex CRF resulting from recombination between subtypes A, G, H and K. It has some unclassified genomic regions as well. Group O or “Outlier” is comprised of strains other than the genotypes found in Group M. It is mainly found in west-central Africa. Group N or “New” was first reported in 1998 in Cameroon and is extremely rare [5].

1.2 AIDS prevalence and general epidemiology

Over the past couple of decades the AIDS epidemic has claimed millions of human lives worldwide. According to the latest statistics released by UNAIDS/WHO in November 2005, there are nearly 40.3 million people estimated to be living with HIV/AIDS in 2005, out of

which 2.3 million are children. About five million people are believed to have been newly infected last year. Further, 2005 saw the death of 3.1 million people due to AIDS. Since 1981, the AIDS pandemic has claimed nearly 25 million human lives worldwide. It is estimated that around 6000 people worldwide get infected with HIV every day. Nearly 95% of those living with AIDS reside in developing countries primarily due to poverty, poor medical and sanitation facilities and a major lack in resources for HIV prevention and treatment. This number is estimated to increase even further in the future. Some of the worst struck areas include Sub-Saharan Africa and Latin America and Central Asia. Subtype B is prevalent in the Americas and Western Europe, whereas A and C are predominant in the African and Asian continents [6]. The following are the routes through which HIV infection can spread:

1. Unprotected vaginal/anal/oral sex with an infected person. The infected individual who harbors the virus in his/her sexual fluids can cause the virus to spread to his/her partner. Oral sex poses a slightly lesser risk, although it is highly risky if a condom is not used, and if one partner has an open cut in his/her mouth.
2. Through sharing of infected needles. This mode of transmission is very common among injection drug users who share needles between themselves.
3. Through infected blood/blood products. This is highly prevalent when unscreened blood/blood products are transfused from an infected individual to a non-infected one.
4. From mother to child. This could occur between an HIV-positive mother and her child either during pregnancy/breast-feeding/child-delivery.
5. Exposure to an infected person's blood. This would happen when sufficient blood from an infected person enters the body of an uninfected one [7].

This global pandemic has currently reached epic proportions. Several international organizations like UNAID, AVERT, WHO, CDC, CIPRA, and NIAID have joined hands in the battle against AIDS. Despite the advent of several effective therapies, AIDS remains near the top as a worldwide killer among infectious diseases. Only lower reparatory infections, which are caused by many different viruses and bacteria, kill more people each year. Obviously, we need to know as much as we can about the virus and it's mode of replication so that potential targets for drug therapy, viral inhibitors and vaccines can be determined.

1.3 Pathogenesis of HIV infection

HIV infects T4 lymphocytes or Helper T cells of the immune system. In a healthy individual, upon infection and antigen presentation, activated T4 cells produce and secrete a multitude of interleukins and cytokines in order to activate macrophages, natural killer cells, CD8+ T cells and other cells of the immune system to fight the infection. There are approximately 800-1200 T4 cells per mm³ of blood in a healthy person. As HIV invades the helper T cells, over the course of several years (typically about 10) this number drops to a staggering 200 or even lower. This leads to a profound immunosuppression, and the individual becomes highly susceptible to several opportunistic infections (OIs), cancers etc. These include a wide range of fungal, bacterial, viral and protozoan infections which are otherwise harmless in a healthy host [2]. Some of the symptoms seen early during an HIV infection include: unexplained weight loss, fatigue, headache, recurrent diarrhea, seborrheic dermatitis etc. During this stage the T4 cells are steadily being depleted. Usually the patient is put on anti-retroviral therapy at this stage. The T4 cell count is an indicator of the level of immune depletion in an AIDS patient. Some of the common OIs seen in AIDS patients when the T4 cell count falls lower than 500 per mm³ of blood are candidiasis, Kaposi's sarcoma,

pulmonary tuberculosis, cryptosporidiosis, Herpes Zoster, pneumococcal pneumonia, and oral hairy leukoplakia. Levels lower than 200 lead to Pnuemocystis carinii pneumonia (PCP, the number one cause of death in AIDS patients), extra-pulmonary tuberculosis, Toxoplasmosis and others. Infections caused due to Cytomegalovirus and Mycobacterium avium are seen when the levels drop further down to 50 [8].

1.4 Anti-retroviral therapy

Unfortunately, to date there is no cure for AIDS or HIV infection. There are anti-retroviral drugs available that reduce the viral load, slow down and possibly stop disease progression. Currently, there are five classes of anti-HIV drugs approved by the FDA. They are:

1. **Nucleoside analog RT inhibitors (NRTI):** These drugs were introduced in 1987. They were the first group of anti-HIV drugs available to treat AIDS patients. These are a class of drugs that interfere with the process of reverse transcription mediated by reverse transcriptase (RT) enzyme. They are popularly known as “nukes” for nucleoside analogues. They affect the synthesis of the growing viral DNA strand by competing with normal nucleotides for binding to RT, then becoming incorporated into the growing chain. The lack of an extendable 3’ hydroxyl group causes the chain to terminate. Some examples of this class of drugs include lamivudine, abacavir, stavudine, zidovudine, and didanosine.
2. **Non-nucleoside analog Reverse Transcriptase inhibitors (NNRTI):** These are a class of drugs that are non-nucleoside analogues that also interfere with the reverse transcription process, although they do so by binding non-competitively to RT. Commonly known as non-nukes or non-nucleosides, these drugs came to the market

around 1997. Some examples of NNRTI drugs include delavirdine, nevirapine and efavirenz.

3. **Protease inhibitors:** These drugs inhibit the viral protease enzymes that function later during viral replication. They were approved a couple of years before the NNRTI drugs in 1995. The protease enzyme functions to cleave precursor proteins into functional peptides. Defective protease enzymes or lack of protease function results in immature, non-infectious virions with aberrant core morphologies. Examples of this class of drugs include ritonavir, saquinavir, nelfinavir and indinavir.

4. **Nucleotide Reverse Transcriptase Inhibitors (NtRTI):** These function like the NRTI drugs except that they are nucleotide analogues. They were approved in October of 2001 by the FDA. Tenofovir is one such example of a nucleotide analogue.

5. **Fusion/entry inhibitors:** This is the fifth and latest class of drugs that have been recently approved in the US and Europe since 2003. Unlike the above three class of drugs that interfere with the viral replication steps that occur after viral entry into the host, this class of drugs acts against HIV prior to its entry into the cell. T-20 is an example of a fusion inhibitor drug that binds to the membrane-spanning domain of the viral glycoprotein polypeptide complex namely gp-41, thereby preventing it from fusing with host cell membrane. Further, entry inhibitors cannot be administered through the oral route, they need to be injected intravenously to prevent them from getting digested.

More than one anti-retroviral drug needs to be taken at a time for the therapy to be effective over a long duration. This type of therapy is commonly referred to as combinational therapy. The low replication fidelity of RT enzyme often causes it to make errors during viral replication. As a result, several divergent strains of the virus are harbored within the patient.

During anti-retroviral treatment, occasionally, a strain could emerge which is resistant to a particular drug. This usually results in failure of the treatment, especially if the patient is not on any other drug. Highly Active Antiretroviral Therapy or HAART is a highly effective combination therapy where three or more anti-retroviral drugs, usually with a protease inhibitor are given to battle drug-resistant mutants [9]. Twenty seven vaccine and fifteen microbicide candidates were currently undergoing human trials worldwide at the beginning of the year 2006. Several others are in experimental stages. Only one candidate vaccine, called ALVAC (a canarypox-based vaccine) in combination with an AIDSVAX (a bivalent, recombinant gp-120 based vaccine) booster is currently undergoing phase III trials in Thailand involving nearly 16,000 volunteers. Results from this trial are expected by the year 2010. Four microbicide candidates have also reached phase III trials. These include SAVVY (a surfactant), PRO2000, Carraguard and cellulose sulfate (adsorption inhibitors). Vaccines employing live-attenuated and whole-killed viruses are currently not undergoing human trials because there is a slight risk that preparations could contain some active virus. Vaccine designs that involve peptide epitopes of the virus and recombinant viral proteins present more hope. Vaccines that employ a combination of designs and/or antigens and those that use bacterial and viral vectors are also promising. The quest for an effective vaccine could go on for several more decades, and it is highly unlikely that an effective one would be engineered even by 2015. However, things look brighter in terms of microbicides and researchers believe that an effective one could be made available by 2012, if sufficient funding and effort is put in [10].

1.5 HIV-1 morphology and genome organization

Mature HIV virions are roughly 80-120 nm in size, spherical particles with many spikes projecting from the outer membrane. A characteristic bullet-shaped core is seen in the center of the virion which encapsidates the genomic RNAs, other RNAs and many viral proteins. Early studies revealed that retroviral particles are roughly composed of about 65% protein, 35% lipids and 1-2% RNA respectively. Genomic RNAs account for nearly two-thirds of the viral RNA, while transfer RNAs, other small RNAs and some host RNAs add up to the remaining one-third of the RNA content. In pure viral preparations, nearly three-quarters of the total protein content are composed of *Gag* structural proteins, followed by the viral envelope glycoproteins, viral enzyme proteins and some host cellular proteins. The HIV genome is roughly 9 kb in size. Retroviruses are diploid with a genome comprised of two positive sense single stranded RNAs that are identical/almost identical to each other. Bender *et al.* extracted RNAs from many retroviral particles and examined them under denaturing conditions. Electron microscopy revealed that they are non-covalently linked at many regions, the strongest being the 5'ends. This occurs through complementarity between many base pairs at the 5'ends. This region is called the Dimer Linkage Site (DLS) and has been demonstrated to be a hot-spot for retroviral recombination in HIV and MuLV (Moloney murine Leukemia Virus). Hence, the genomic RNA is said to be dimeric in nature. This dimeric nature of the genome is primarily responsible for the high rate of recombination in HIV, since sequences from both the RNAs could be copied into proviral DNA during reverse transcription. Although virions contain two genome copies, only a single dsDNA is produced. The presence of a second genomic RNA may act as a salvage pathway during reverse transcription when broken or damaged regions in one of the RNAs are encountered

reverse transcription can switch to the second. Two genomes also allow for extensive recombination and therefore greatly contribute to genetic diversity [1].

Figure 1-2 shows the organization of the HIV genomic RNA. It is approximately 9,200 nucleotides in length. The viral RNA shares many similarities to an mRNA. There is a 5'-5' linkage between the first encoded nucleotide of the genomic RNA and a methylated GDP residue which serves as a cap at the 5' end. Following the methylated capped 5' end is the R (Repeat) region. The U5 (Unique sequence at the 5' end) region and the PBS (Primer Binding Site) sequences are present after this. This is followed by the ORFs (Open Reading Frame) of the nine genes of HIV. These nine genes code for nine different proteins. Following this are the PPT (polypurine tract), U3 (Unique sequence at 3' end) and R regions. The 3' end of the genome is polyadenylated. It bears a poly (A) tail composed of nearly 50-200 non-coding A residues. The virus makes a complete dsDNA copy of the RNA genome. This DNA is referred to as the proviral DNA after it inserts in the host genome. A schematic of the proviral DNA is also shown in Fig. 1-2. It can be noted that the proviral DNA contains two LTR regions at both ends of the genome. These regions are called Long Terminal Repeats and are found only in the proviral DNA. They are made as a result of duplication of the U5 and U3 regions during reverse transcription. LTRs are very vital for the obligatory second strand transfer event that occurs during reverse transcription to complete proviral DNA synthesis. Further, they also contain the promoter sequence which is mandatory for the transcription event mediated by the host RNA polymerase II enzyme. The 3' LTRs contain some polyadenylation sites which are important for transcription as well [1]. Table 1-1 list the nine genes of HIV and summarizes their function in the viral life cycle.

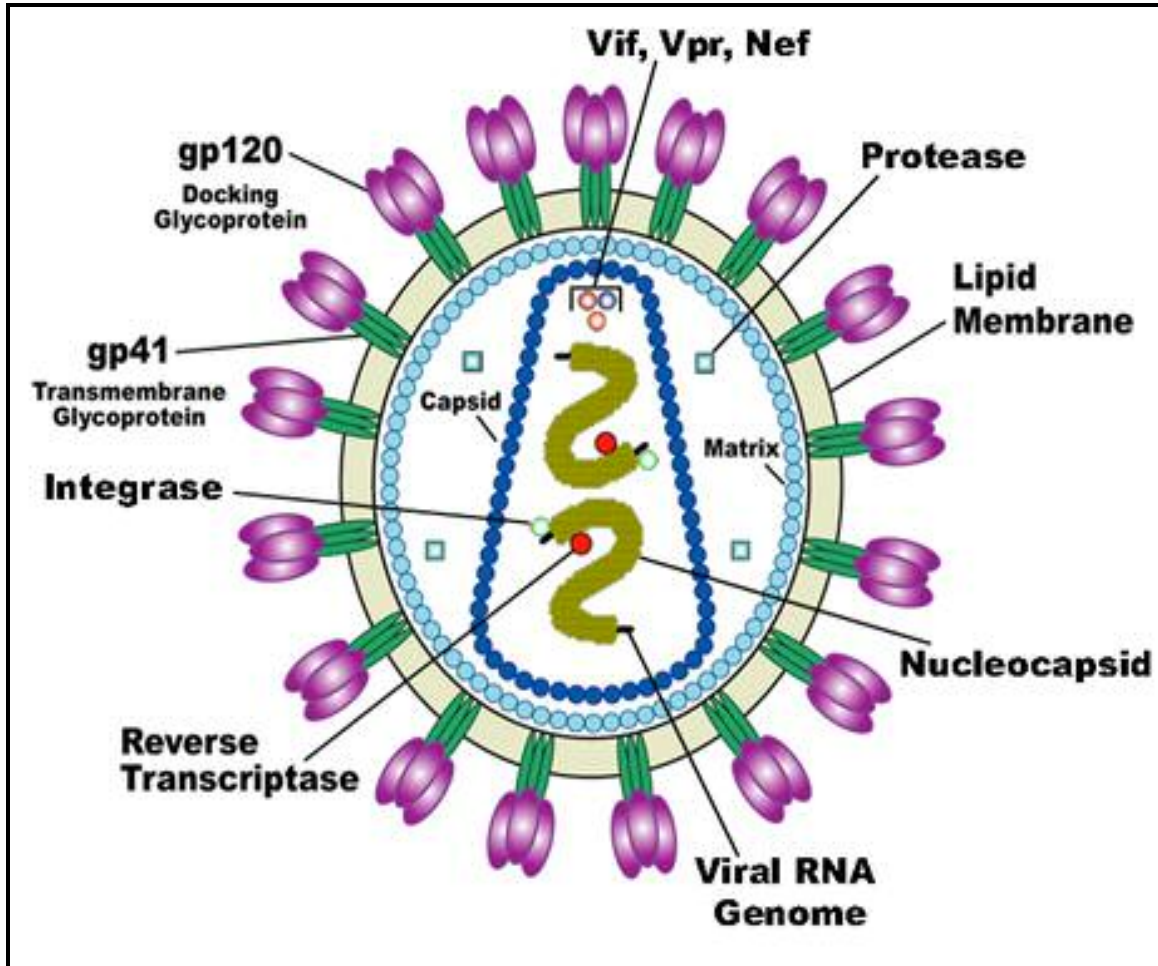


Figure 1-1: Schematic of mature HIV-1 virion

Shown above is a schematic diagram of a mature HIV-1 virion. The viral RNA genome consisting of two single-stranded, positive sense RNA strands coated with the nucleocapsid proteins are shown. The viral capsid which encloses the genome is shown in blue. The capsid also encloses the viral enzymes namely reverse transcriptase (RT), integrase (IN) and accessory proteins namely vpr, nef and vif. The capsid made up of capsid protein (CA) is surrounded by the matrix (MA). The viral protease enzyme (PR) is found within the matrix. The matrix is enclosed within the viral envelope. The envelope glycoproteins namely gp120 (SU-surface unit) and gp41 (TM-transmembrane protein) are also shown. Figure obtained from “AIDS Fact sheets and Brochures – How HIV Causes AIDS” maintained by National Institute of Allergy and Infectious Diseases (<http://www.niaid.nih.gov/factsheets/howhiv.htm>).

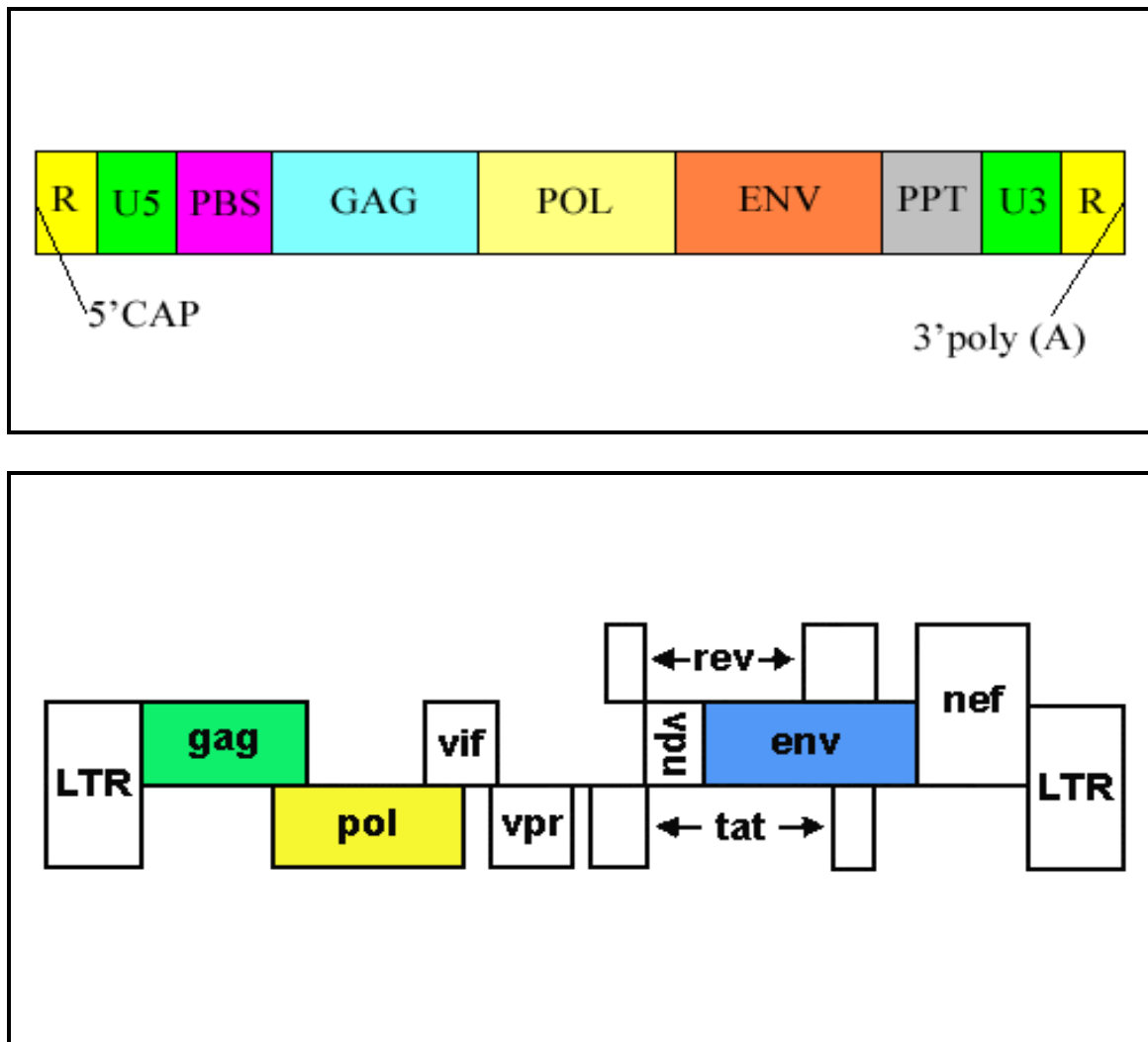


Figure 1-2: HIV-1 RNA genome and proviral DNA

Shown above are the viral RNA genome and proviral DNA. The RNA genome contains a methylated GDP cap at the 5' end. This is followed by the R (repeat), U5 (unique region at 5' end) and PBS (primer binding site) regions. The t-RNA primer binds to the PBS to initiate reverse transcription. This is followed by the *gag*, *pol* and *env* coding regions. The PPT, U3 and R regions are found after the coding regions. PPT is the polypurine tract which acts as a primer for plus sense strong stop DNA (+ssDNA) synthesis. U3 and R regions, respectively, are unique and repeat regions at the 3' end. The HIV proviral DNA is also shown. The nine genes of HIV-1 and their respective locations are also depicted. In some cases (Tat and Rev) the proteins are produced from non-contiguous regions of RNA that are spliced together. Note the presence of LTR (long terminal repeats) regions on either ends of the genome. LTRs are very crucial in order to complete second transfer step during reverse transcription. Figure obtained from "HIV/AIDS Resources on the Internet" maintained by The University of Zambia, School of Medicine (<http://www.medguide.org.zm/aids/HIVgenom.gif>).

Table 1-1
The nine genes of HIV-1

Gene name(s) (5' to 3'end)	Proteins encoded (suffix indicates molecular mass in kDa)	Function in viral life cycle
<i>gag</i> (group specific antigen)	p17 (matrix) p24 (capsid) p7 (NC-nucleocapsid)	nuclear localization of preintegration complex, protection, assembly protection and assembly protection, assembly, RNA packaging, integration, t-RNA binding, nucleic acid chaperone, enhances RT processivity, viral recombination
<i>pol</i> (polymerase)	p10 (protease) p31 (integrase) p66/51 (RT-reverse transcriptase)	cleavage of precursor polyproteins into functional peptides integration of proviral DNA makes a DNA copy of viral RNA genome
<i>env</i> (envelope)	gp120 (SU) gp41 (TM)	surface envelope glycoprotein binds to CD4 antigens on helper T cells. transmembrane protein which mediates fusion with host cell membrane
<i>tat</i> (trans-activator of transcription)	p14	activates viral transcription
<i>rev</i> (regulator of virion protein expression)	p13/19	splicing of viral RNA, nuclear export of RNA
<i>vpr</i> (viral protein R)	p15	CD4 degradation, enhances infectivity of virions, interferes with T cell signal transduction
<i>vpu</i> (viral protein U)	p16	release of virions, CD4 degradation
<i>vif</i> (virion infectivity factor)	p23	enhances virion infectivity
<i>nef</i> (negative regulatory factor)	p27	enhances virion infectivity, interferes with activation of T-cells, regulates viral replication

1.6 Viral life cycle

1.6.1 Attachment and entry: The HIV virion interacts with the host cell membrane by means of the globular surface glycoprotein (SU) gp120. The CD4 antigen present on the surface of phagocytic cells like macrophages, monocytes, helper-T cells and others serves as the primary receptor for recognition by gp120. In addition to CD4, the virus needs another co-receptor for attachment and entry. This co-receptor is of two kinds namely CXCR4 and CCR5. FUSIN or CXCR4 is an α -chemokine receptor found on T-cells. Hence the syncytia inducing (SI) strains of HIV-1 which use the FUSIN co-receptor for entry are named T-tropic. Whereas the non-syncytia inducing (NSI) strains of the virus that use the CCR5 (a β -chemokine receptor) receptor found on macrophages are named M-tropic. Viral strains that use only the CCR5 co-receptor for attachment are called R5 while those that use only the CXCR4 co-receptor are termed X4. This interaction exposes the V3 loop of gp120. The conformational change in gp120 exposes a fusion peptide in gp41, thereby enabling it to mediate fusion with the host cell envelope. Once the viral envelope fuses, the capsid containing the genome and other viral proteins is released into the cytoplasm [11].

1.6.2 Viral replication: Viral replication occurs within the ribonucleoprotein complex which encloses RNA, RT, NC and other viral proteins. The activity of RT was extensively studied by Howard Temin and David Baltimore by monitoring replication intermediates in purified virions. In 1970 their independent discoveries of an RNA-dependant DNA polymerase or reverse transcriptase was reported and these scientists later shared a Nobel Prize for the discovery. Reverse transcriptase (RT) is a multi-functional enzyme that possesses RNA-dependant DNA polymerase activity, DNA-dependant DNA polymerase activity and RNase H activity. The retroviral nucleocapsid protein (NC) also has a multitude

of functions during viral replication. The virus generates a complete dsDNA copy of its RNA genome by reverse transcription. Figure 1-3 [12] shows a schematic diagram of the process of reverse transcription. The viral genomic RNA is denoted in black; minus-strand and plus-strand DNAs are denoted in orange and red respectively. The following are the various steps that occur during reverse transcription:

1. The partially unwound t-RNA primer is annealed to the PBS at the 5' end of the RNA genome by means of an 18 base complementary region. RT initiates minus strand DNA synthesis using the 3' end of this t-RNA primer. Synthesis proceeds till the 5' end of the RNA is reached generating a short DNA intermediate termed –sss DNA (minus-strand strong stop DNA).
2. RNase H activity of RT degrades the RNA from the RNA/-sss DNA hybrid. The complementarity in the R regions between the 3' ends of –sss DNA and viral RNA allows the first strand transfer/jump of the –sss DNA to the 3' end of the RNA.
3. Extension of the minus strand DNA resumes and continues to the PBS region at the 5' end of the RNA genome. This is accompanied by RT mediated RNase H degradation of the template RNA.
4. PPT or polypurine tract remains resistant to the RNase H degradation of RT and serves as the primer for +sss (plus-strand strong stop) DNA synthesis. Plus-strand DNA synthesis proceeds till the PBS region on the t-RNA is reached. A modified base at the 19th position from 3' end of the tRNA results in termination of synthesis after the first 18 bases are copied. Subsequently, RT degrades the t-RNA primer and the PPT region.
5. The complementarity between PBS regions of the minus-strand DNA and the +sss DNA mediates the second transfer event.

6. The minus-strand DNA/+sss DNA duplex then circularizes in order to serve as templates for one another.

7. Synthesis of minus-strand and plus-strand DNA is then completed, thereby generating a complete blunt-ended DNA copy of the viral RNA genome. Duplication of the U5 and U3 sequences generates the LTRs (long terminal repeats) on either ends of the proviral DNA [12].

1.6.3 Integration, assembly and release: The viral DNA then integrates randomly into the host chromosomal DNA as part of a pre-integration complex. Although integration is random and non-uniform, highly bent regions such as nucleosomes are preferred target sites. The integrated DNA is referred to as the provirus. The virus then uses host cell machinery to generate viral RNA and proteins. Assembly of viral RNA and proteins then begins near the plasma membrane of the host cell (See Fig. 1-4). HIV packages two RNA genomes along with viral proteins and host tRNAs into each virion. The viral glycoproteins then associate with the inner surface of the host cell's plasma membrane and begin to bud from the cell. The virion is immature initially. Maturation occurs either in the budding virion or after budding is completed. Maturation is brought about by viral protease enzymes that cleave precursor polyproteins into functional peptides and enzymes. This results in the formation of a mature HIV virion capable of infecting other cells [13], [14], [15].

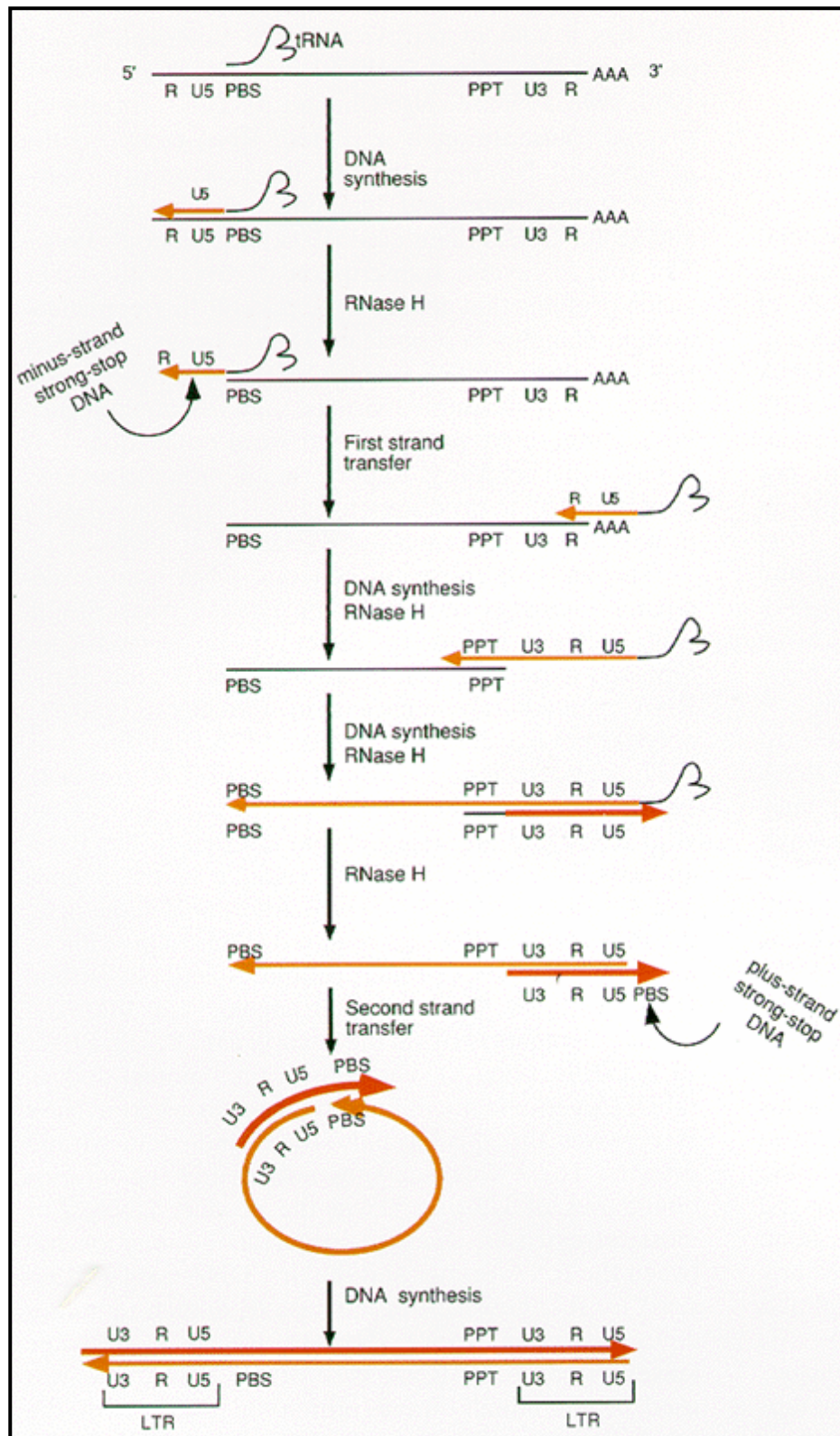


Figure 1-3: Schematic of Reverse transcription: Shown is the schematic model of reverse transcription illustrating the various events described above.

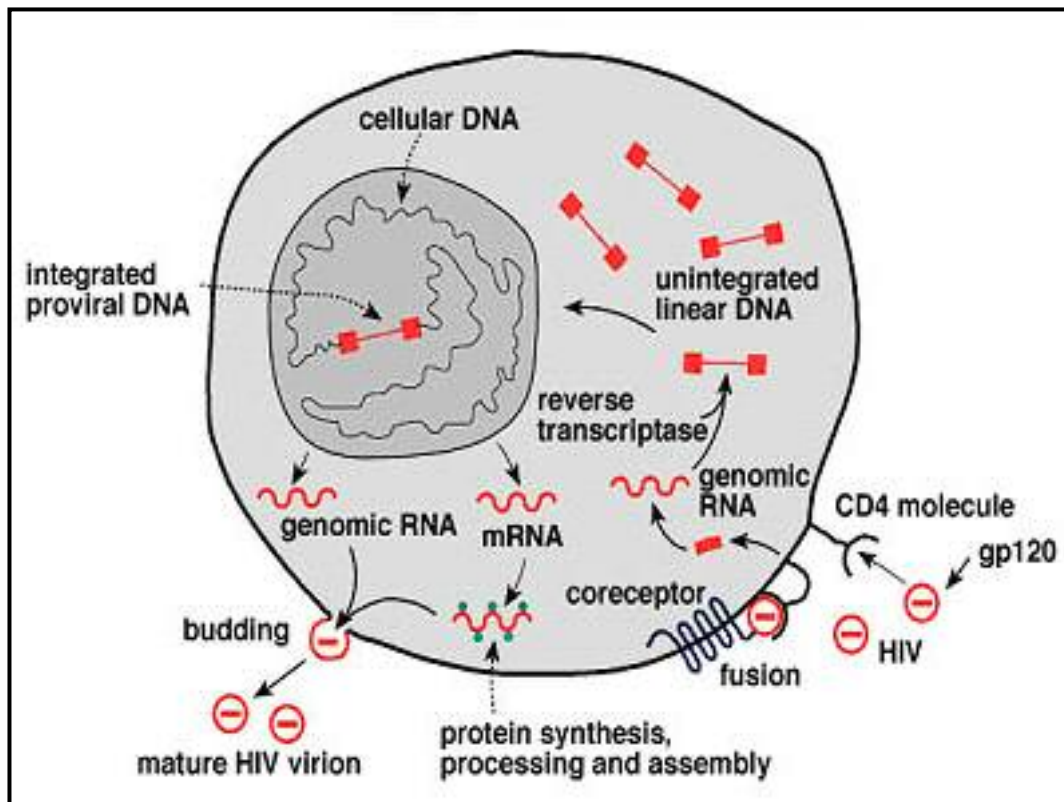


Figure 1-4: Viral life cycle

Shown above is a schematic diagram illustrating the various events occurring during the HIV viral life cycle. The virus first attaches to the host cell by means its surface glycoprotein gp120. It then fuses with the host cell membrane and releases a nucleoprotein complex into the cytoplasm. A complete dsDNA copy of the viral genome is made via reverse transcription. The dsDNA then randomly integrates into the host chromosomal DNA as a provirus. Synthesis of viral RNA and proteins occurs using the host cell machinery. These components then assemble near the host cell membrane where packaging occurs. The viral glycoproteins then fuse with the host cell membrane and immature virions bud off. Viral proteases complete maturation through proteolytic cleavage. Figure obtained from “AIDS Fact sheets and Brochures – How HIV Causes AIDS” maintained by National Institute of Allergy and Infectious Diseases (<http://www.niaid.nih.gov/factsheets/howhiv.htm>).

1.7 Reverse transcriptase

The characteristic feature that distinguishes retroviruses from other viruses is the conversion of their RNA genome into a DNA intermediate which integrates into the host genome. The enzyme that brings about this “retro” or reverse flow of genetic information contrary to the central dogma of genetics is reverse transcriptase (RT). HIV RT was independently discovered by Howard Temin and David Baltimore in the year 1970 and was heralded as one of the most significant contributions to the field of retrovirology. RT possesses three independent activities namely: RNA-dependant DNA polymerase, DNA-dependant DNA polymerase and RNase H activity. HIV-RT is active as a heterodimeric protein consisting of two subunits with similar amino termini. The larger subunit is called p66 and the smaller one is called p51. The suffix denotes their molecular weight in kDa. The p66 subunit is 560 amino acids in length. The p51 subunit is produced as a result of proteolytic cleavage from the carboxyl terminus of p66. Hence, it consists of the first 440 amino acids of p66. The p66 subunit folds into two domains, the polymerase and RNase H domain. Since the p51 subunit corresponds to the first 440 amino acids of p66, it contains only the polymerase domain and lacks the RNase H domain (from the carboxyl terminus) [16]. The exact role of p51 is still not completely clear. Experiments carried out with HIV-1 specific inhibitors and certain thymidine triphosphate templates suggest that the polymerase activity lies exclusively within the p66 subunit. Hostomsky *et al.* have shown that mutations disrupting the polymerase active site in the p66/p51 heterodimer of RT can be tolerated by p51 but not by p66. Le Grice *et al.* have conducted mutational studies in which the p51 subunit was inactivated by mutations and found that all the activities of RT were carried out by the p66 subunit. This suggests that p51 is not directly involved in the polymerase and

RNase H activities of RT [17]. Figure 1-5 shows a color-coded schematic ribbon diagram of RT. RT folds into the characteristic open right handed conformation similar to many other prokaryotic and eukaryotic DNA polymerases. The three domains namely fingers, palm and thumb that give it the shape of an open right hand are shown in blue, red and green respectively. The connection domain between the two subunits is shown in yellow. The active site of the enzyme is found in the palm domain. The template is held in the groove of the hand. The thumb touches the minor groove of the template. The fingers enclose the primer portion of the primer-template hybrid. Larder *et al.* have shown through site-directed mutagenesis that the polymerase domain active site contains three aspartic acid residues: Asp110, Asp185 and Asp186 [18]. This lies in the N-terminal domain as shown in the linear schematic diagram in Fig 1-6. The RNase H domain lies in the C-terminus and is shown in gold. The four active site residues of this domain are Asp443, Glu478, Asp498 and Asp549 [19], [20]. Studies have shown that the distance between these two active sites is around 18/19 nucleotides. This is represented linearly as well (See Fig 1-6) [21]. The highly negatively charged catalytic core formed due to the active site amino acids enables RT to coordinate Mg^{2+} . Processivity of an enzyme refers to the average number of nucleotides that it incorporates into the growing nascent strand in a single binding event. HIV-1 RT's processivity is approximately 100 which is low for a replicative polymerase. It does not possess a 3' to 5' proofreading/exonuclease activity. Further, it also has low fidelity with an estimated error rate of 10^{-4} to 10^{-5} per base incorporated. The low fidelity coupled with a high recombination rate is a major driving force for generating genetic diversity that is prevalent with HIV [22], [23].

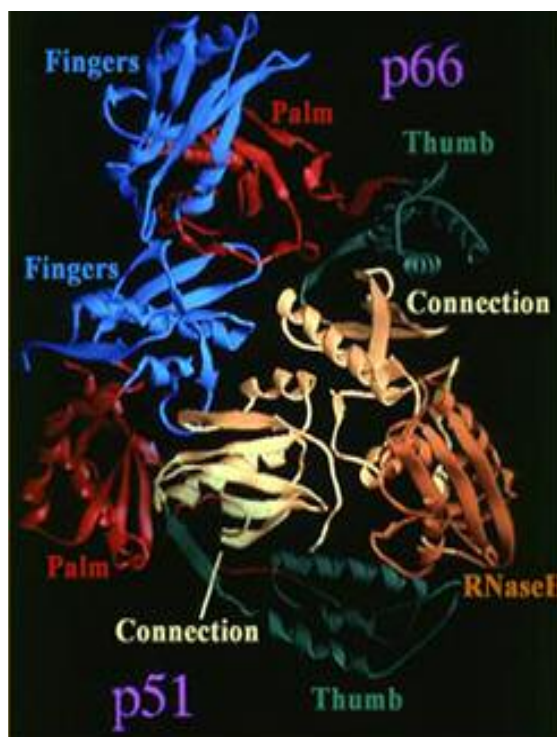


Figure 1-5: HIV-1 RT

Shown is a schematic ribbon diagram of HIV-1 reverse transcriptase (RT) enzyme. It is a heterodimeric enzyme that contains p51 and p66 subunits. The p66 subunit contains polymerase and RNase H domains, while the p51 lacks the RNase H domain. The polymerase domain bears the characteristic open right hand conformation with the fingers (blue), palm (red) and thumb (green) domains. The connection domain linking the two subunits is shown in yellow, while the RNase H domain is shown in gold. Figure obtained from “Overview of RT Structure” maintained by HIV Drug Resistance Program, National Cancer Institute (http://www.retrovirus.info/rt/overview_f1.html).

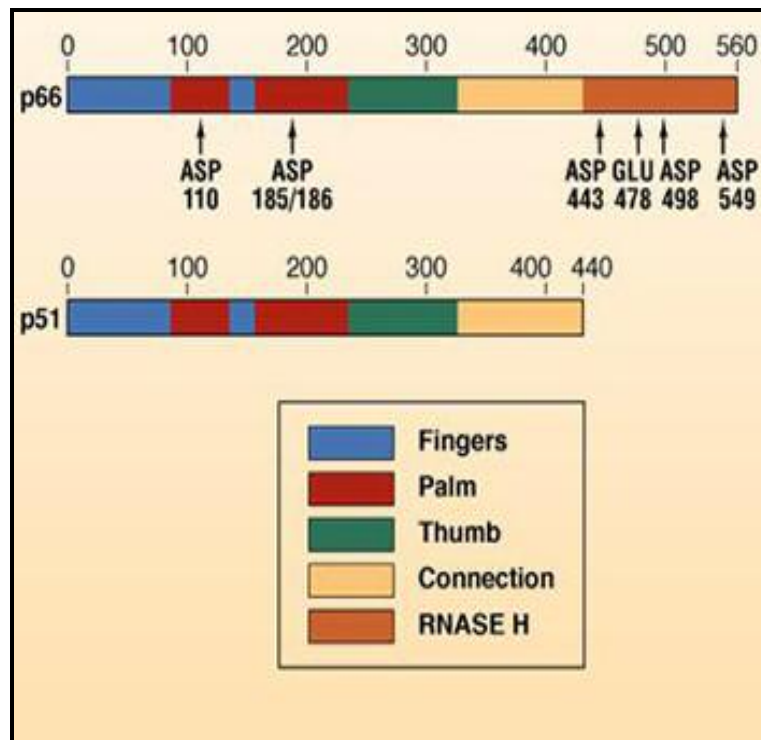


Figure 1-6: Linear diagram of HIV-1 RT

Shown is a linear diagram representing the two RT subunits, p66 and p51. The p66 subunit contains both polymerase and RNase H domains and is 560 amino acids in length. The polymerase active site catalytic residues (Asp110, Asp185 and Asp186) are indicated. The RNase H active site catalytic residues (Asp443, Glu478, Asp498 and Asp549) are also shown. The p51 subunit which is produced as a result of proteolytic cleavage from the C-terminal of p66 is shorter and lacks the RNase H domain. Hence it contains 440 amino acids. The color coordination between the ribbon diagram shown previously (See Fig. 1-5) and the linear representation is the same. Figure obtained from “Overview of RT Structure” maintained by HIV Drug Resistance Program, National Cancer Institute (http://www.retrovirus.info/rt/overview_f1.html).

1.8 Recombination and genetic diversity in HIV

HIV and retroviruses in general possess one of the highest recombinogenic potentials among viruses. This is well documented by the huge number of recombinant genetic variants that are found within HIV. As was noted earlier in section 1.1, there are several established subtypes within the HIV-1 species. These most likely represent divergent evolution from a common progenitor. In recent years co-infection of individual hosts with different subtypes has resulted in new circulating recombinant forms (CRFs) that have become established in various regions of the world. The CRFs contain genome regions derived from two or more subtypes. The well established CRFs originate diverse recombinant virus forms that are also found in many regions. The term intersubtype recombinant or ISR is used to refer to these genetically diverse recombinants. For example, any recombinant containing genome regions from an A and E virus is an A/E ISR while a specific A/E CRF such as CRF01_A/E is the most prevalent virus infecting people in Thailand. The huge genetic diversity of HIV-1 has caused the HIV population to be referred to as a quasispecies which is defined as a “dynamic distribution of nonidentical but related viral replicons” [24].

HIV continually mutates and generates genetic variants. It circumvents drug therapy by producing drug-resistant mutants. Several factors have been found to be responsible for the huge genetic diversity of HIV. They are:

1. Low fidelity of RT enzyme and lack of 3'-5' proofreading/exonuclease activity.
2. Recombination (This will be discussed in detail shortly).
3. Presence of diploid genome provides a greater chance for recombination during reverse transcription.

4. Possible low fidelity of host RNA Pol II enzyme that is involved in the virus life cycle.
5. High rate of viral replication: As many as 10^{10} new virions are produced per day in an infected individual. This increases the potential for a large number of new mutant strains.

Further, HIV also evades the immune response generated by the host by employing several pathways. Some of them include:

1. Continual depletion of immune cells leading to progressive immune deficiency.
2. Cell to cell adhesion and syncytia formation causing infection to spread without a virion being exposed to the blood stream and immune system.
3. Lack of viral proteins in the plasma membrane during macrophage infection curtails immune stimulation and recognition.
4. Long term latency in some infected immune cells serving as a reservoir for future infections.

Recombination is one of the main factors responsible for generating genetic diversity within HIV. The diploid nature of the retroviral genome suggests that recombination may be a crucial function necessary for the virus to replicate and evolve. HIV packages two RNAs within each virion. The RNA genomes could be identical or almost identical. This would likely be the case in cells infected by a single virus as the genomes would be derived from one provirus. In these cases the fidelity of host RNA pol II would be the major source of diversity. Although the exact fidelity of this enzyme is unknown, results suggest it has considerably higher fidelity than RT [25]. Recombination between identical genomes would not produce genetic diversity. In some cases cells may be infected by more than one virus

and in these cases multiple proviruses can be made. Different genomes can become packaged in the same virion in these cells and the impact of recombination is greater. For simple retroviruses such as spleen necrosis virus (SNV) and MuLV, homologous recombination was shown to occur in nearly 4% (the rate is much greater in HIV as noted below) of DNAs in progeny virions during one replication cycle when genetic markers spanning 1 kb were used. This corresponds to about a 40% chance of recombination per replication round for a 10,000 base genome. The rate of non-homologous recombination is significantly lower; only 1/100 to 1/1000 [26]. Recombination in HIV occurs through a process called strand transfer or template switching (also called strand jumping). Strand transfer refers to the process by which nascent DNA being synthesized on one RNA strand (referred as donor) falls off and jumps onto homologous regions on either the same RNA or on another RNA strand (referred as acceptor) where synthesis is completed. If the transfer occurs to the acceptor template, the resultant DNA would be a chimera between the parent and acceptor RNA templates. As described earlier in section 1.6.2, the process of viral replication requires two mandatory strand transfer reactions (also called strong stop DNA transfers) in order to complete reverse transcription. The first strand transfer occurs when the nascent –sss DNA (minus-strand strong-stop DNA) jumps and anneals to the 3'end of genomic RNA. This is mediated through complementarity between the R or repeat regions. The second strand jump occurs when the nascent +sssDNA anneals to the complementary PBS region found in the minus-strand DNA resulting in circularizing of the genome and completion of plus-strand DNA synthesis. These strand transfer events occur at the end/termini of the genome. Hence, they are termed as end/terminal strand transfers. Besides end strand transfer, the virus can also undergo numerous internal strand transfers which occur randomly anywhere throughout the

genome [27], [28]. These events can occur anytime during the minus/plus strand DNA synthesis. Zhuang *et al.* have shown that the average rate of recombination in HIV is about 2.8 crossovers per cycle of viral replication [29]. Cell infection experiments conducted by Levy *et al.* in T-lymphocytes showed a higher recombination rate of about 10 cross-overs per replication cycle. Nearly thirty crossover events were reported to have occurred during infection of macrophages. The latter figure represents approximately one recombination event for every 350 nucleotides synthesized [30]. In one report it was demonstrated that HIV-1 recombines ten to twenty times more frequently than MuLV. The reason for the difference is unclear, although HIV is considered a “complex” retrovirus meaning it produces several regulatory proteins (Tat and Rev) for example, while MuLV is a “simple” retrovirus. Virion morphology is also quite different for these viruses [28].

Currently, there are two models that explain the mechanism of retroviral recombination: copy choice and strand-displacement assimilation. The copy choice model explains strand transfers that occur during minus-strand DNA synthesis. According to this model, the nascent DNA switches from one RNA template to another during minus-strand synthesis. This model is a modified version of the original “forced copy-choice” model which postulates that breaks in the RNA genome force the nascent DNA to switch templates in order to complete synthesis [31]. Subsequently it was shown that breaks were not required for strand transfer to occur. Experimental evidence provided by Hu and Temin showed that while γ -radiation, which induced genome breaks, reduced the recovery of viral RNA, it did not increase the frequency of viral recombination [32]. Efficient recombination was reported in experiments where unbroken RNA templates were used [33]. DeStefano *et al.* have shown that efficient template switching occurs between internal regions of RNA templates [34]. In regions of

unbroken RNA, other factors such as the presence of secondary structures could cause stalling of RT, thereby inducing strand transfer. The strand-displacement assimilation model explains strand transfers that occur during plus-strand DNA synthesis. This model was proposed based on observations made on H-branched structures that are classified as intermediates during recombination [35]. Though plus-strand synthesis is initiated predominantly at the polypurine tract, several reports have demonstrated that in some retroviruses, it could be initiated discontinuously throughout the genome at many points [36]. According to this model, at any given time, there are many nascent plus-strand DNA fragments bound to the minus-strand DNA template. The 5'ends of these nascent plus-strand DNAs tend to get displaced by the 3'ends of the preceding plus-strand DNAs. This results in the displacement of such nascent plus-strand DNAs. These displaced fragments could either base pair with the other minus-strand DNA template or randomly integrate themselves between the remaining plus-strand DNA fragments bound to the parent minus-strand DNA template. Either pathway would result in genetic recombination. It should be noted that this model requires the presence of two minus-strand DNA copies of the homologous region where recombination could potentially occur. Results indicate that majority of recombination in retroviruses occurs during minus strand DNA synthesis, probably by copy choice.

1.9 Nucleocapsid protein

The nucleocapsid protein (NC) of HIV-1 is a small (55 amino acid) protein that coats the genomic RNA in the mature virion. The presence of fifteen arginine and lysine residues makes it highly basic. It has a net positive charge of +13 and a pI between 10.0 and 11.0. Nearly 2000 molecules of NC are found within the virion core [37]. The *gag* gene of retroviruses encodes the *Gag* precursor polyprotein. Retroviral internal structural proteins are

derived from *Gag*. It is involved in copackaging of two unspliced RNA genomes and hence plays an important role in viral assembly. After assembly, the immature virion buds off. Maturation could occur either at the time of budding or after budding. It involves cleavage of precursor polyproteins by the viral enzyme protease to yield smaller functional proteins. NC is produced as a result of this proteolytic cleavage [38], [39]. All known orthoretroviruses (except spumaviruses) contain one or two -Cys-X₂-Cys-X₄-His-X₄-Cys- (CCHC) zinc finger motifs as part of NC. Figure 1-7 shows a schematic diagram of HIV-1 NC. In solution, HIV NC contains two rigid fourteen residue CX₂CX₄HX₄C (where X can be any amino acid) motifs, where three cysteines and one histidine coordinate one zinc ion. The strong coordination between the zinc ions and the fingers exists in virions and under *in vitro* conditions as well. The presence of a cluster of arginine or lysine residues near CCHC motifs is typical, and this is seen in NC as well. The two zinc fingers of NC are connected by a highly basic seven amino acid linker (RAPRKKG) [37], [40], [41]. The proline residue found in the linker has been shown to slightly bend the linker such that proximity between the two fingers is increased [42]. Changes that affect zinc binding or deletion of zinc fingers are not tolerated resulting in defective RNA packaging or absence of genomic RNA in virions. Thus, zinc fingers are potential targets for anti-viral agents that eject zinc. The two zinc fingers of NC are flanked by the N and C terminal tails. The tails have been shown to be highly flexible and possess no predicted structure by NMR [42]. However circular-dichroism studies have predicted that N and C terminal regions could form helices and that there is clustering of the N-terminal basic amino acid residues on one side of the predicted helix [43]. The 120 nt (nucleotide) long viral packaging signal called the ψ site is found at the 5' end of unspliced RNAs. NMR studies have revealed interactions between NC and four RNA stem loops (SL1-

SL4) that constitute the ψ packaging signal. These interactions are crucial for viral assembly and packaging. More specifically, the following interactions were found between NC and SL3. The lysine (Lys3) and arginine (Arg10) residues at the N-terminal zinc finger form a 3_{10} helix. This helix binds with the major groove of SL3. There is a hydrophobic pocket within finger one which is formed of Val13, Phe16, Ile24 and Ala25. Hydrogen bonding occurs between the guanosine residues of the RNA and amide atoms of Phe16 and Ala25. Another hydrophobic cleft consisting of Trp37, Gln45 and Met46 is found in finger two. This is formed due to hydrophobic interactions between the side chains of these residues. Electrostatic interactions are found between Arg32 and negatively charged phosphodiester backbone of the RNA. There are few differences between NC interactions with SL2 and SL3. The 3_{10} helix binds to the minor groove of SL2 and predominantly, electrostatic interactions occur between NC and SL2 [44].

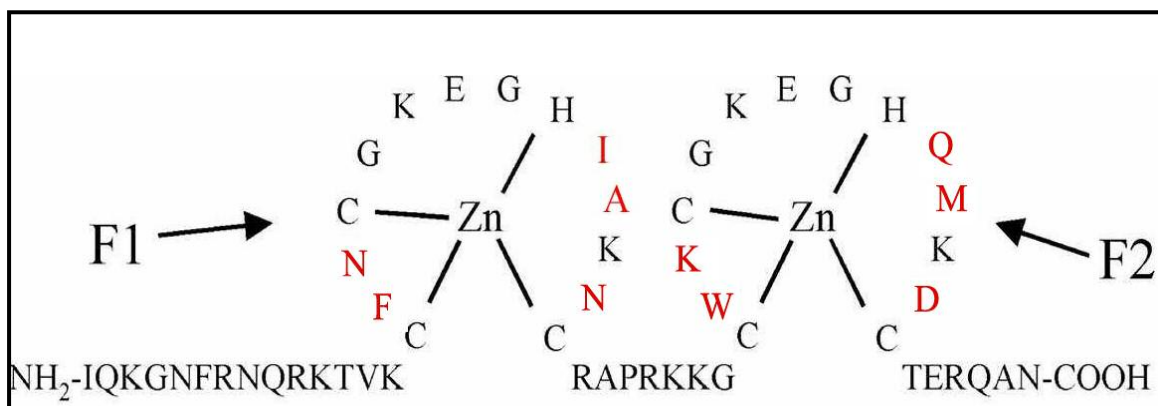


Figure 1-7: Schematic of HIV-1 pNL4-3 nucleocapsid protein

Shown is a schematic diagram of HIV-1 nucleocapsid protein (NC). It is 55 amino acids in length and contains 15 arginine and lysine residues. NC contains two CCHC zinc fingers motifs. The fingers are numbered from the NH₂ terminus and are termed as F1 and F2. The three cysteines and one histidine residue in each finger serve to coordinate one zinc ion. The amino acid differences between the two fingers are indicated in red.

NC is a multifunctional protein and contributes significantly to several important steps in the viral replication cycle [37], [45], [46], [47], [48]. Thus, NC is a potential target for NC-inhibitor drugs [49], [50], [51] and vaccine development [52]. NC coats the viral RNA, though not completely at saturating levels, thereby protecting it from nucleases [53]. Although NC binds to both RNA and DNA, it exhibits preferential binding to single-stranded nucleic acids. NC has been shown to bind to RNA in the following order of preference: retroviral RNA > mRNA > rRNA > poly r(A) [37], [54]. NC is involved in several steps during viral replication. Reverse transcription is initiated by the host t-RNA^{Lys,3} primer that remains bound to the PBS region of the viral genome. NC enhances the unwinding of the t-RNA primer [47], [55] and stimulates its annealing to the PBS [56], [57], [58]. Similar properties were also observed in *in-vitro* experiments conducted with MuLV NC mutants, where basic residues of this NC were shown to enhance t-RNA annealing [59], [60]. NC has been shown to modestly enhance the processivity of RT during reverse transcription [61], [62], [63]. *In vitro* studies have revealed that NC stimulates and modulates the RNase H activity of RT [64], [65]. NC has been shown to stimulate both minus and plus-strand strong-stop DNA transfers and viral recombination in general [66], [67], [68], [69], [70], [71], [72]. More specifically, the levels of transfer efficiency in the first strand transfer event increase by nearly 60% upon addition of NC [41]. Wu *et al.* have monitored plus-strand strong-stop DNA transfer in detail both in endogenous and reconstituted systems. They have demonstrated that NC facilitates removal of the t-RNA primer to complete annealing of PBS sequences during the second strand transfer reaction [73]. In particular, Guo *et al.* have shown that the invariant zinc fingers of NC are important for complete t-RNA removal [46].

Results

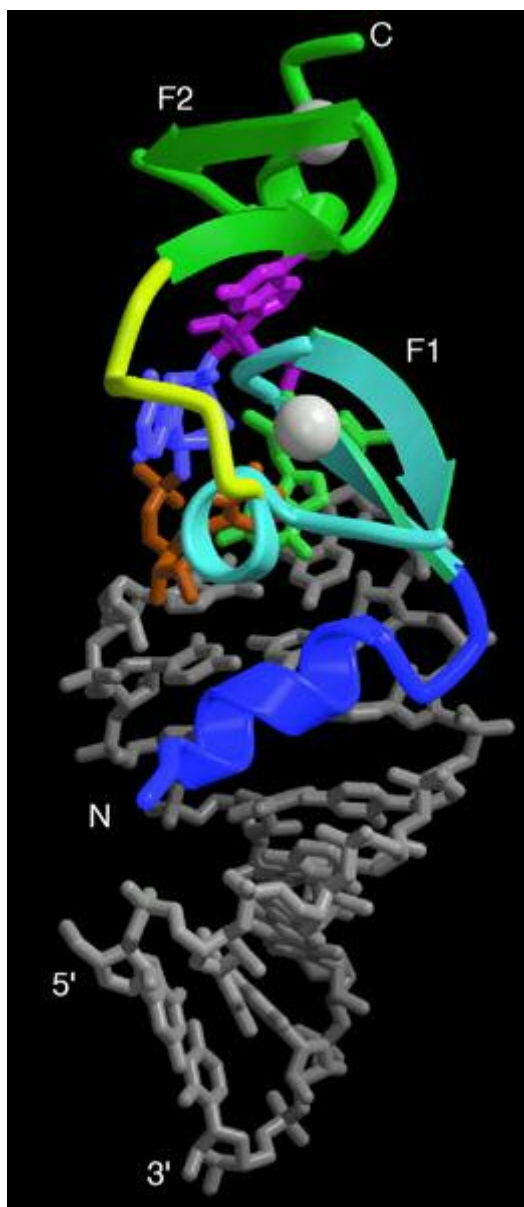


Figure 1-8: HIV-1 NC-SL3 ψ RNA complex

Shown is the ribbon diagram of HIV-1 NC SL3 ψ complex. The color coding is as follows: RNA-gray; first zinc finger-turquoise; linker-yellow; second zinc finger-green; zinc atoms-white spheres; 3_{10} helix-blue; Colored nucleobases in RNA: A⁸-blue, G⁹-dark orange, G⁷-violet, G⁶-light green. Figure obtained from De Guzman *et al* [44].

from many reports indicate that NC also reduces RT pausing and promotes internal strand transfers throughout the genome [45], [63], [67], [71]. Studies show that NC is possibly involved in the integration and protection of the newly synthesized viral genome [74], [75], [53], [76]. As described previously, specific interactions between the NC part of the *Gag* precursor polyprotein and the ψ packaging signal have been observed [77]. Thus, NC is also involved in recognition and packaging of the genome [78], [79], [80], [81]. Upon budding from the host cell, the virion is immature. The two almost identical RNA genomes are held together loosely by weak interactions at the 5'ends. During maturation, the RNA dimer undergoes conformational changes which enable it to become more compact and thermodynamically stable. NC has been shown to enhance the formation of a dimer which has the maximum number of base pairs. This enables the genome to overcome the unfavorable loop-loop dimer conformation and assume a more compact and stable form [82]. NC has been shown to play a role in genomic maturation by promoting dimerization between the two RNAs [83], [84], [85]. Thus, it can be clearly seen that NC is involved in several steps of the viral replication cycle.

NC is a classic example of a nucleic acid chaperone protein [37], [41]. Chaperones are proteins that help nucleic acids to attain the optimal thermodynamically stable conformation. These proteins transiently break base pairs in nucleic acids thereby unfolding them. Subsequently they catalyze the re-pairing or annealing of free bases into an optimal conformation that has the maximum number of base pairs. This enables nucleic acids to overcome suboptimal conformations and unfavorable kinetic traps [86].

One report has demonstrated that NC inhibits annealing of very short, unstructured, complementary sequences [87]. NC presumably accelerates annealing of non-structured

complements provided that they are long enough to form a stable hybrid capable of overcoming NC's helix-destabilizing activity. The observed inhibition could be due to short regions of complementarity. This is further supported by results showing that NC can also inhibit binding of longer complementary regions when non-complementary bases that weaken the hybrid are inserted along the region [88]. Several reports demonstrate that NC enhances annealing even on non-retroviral nucleic acid templates [89], [63], [90]. Thus, the chaperone activity of NC has been well documented by several studies conducted on different types of nucleic acids.

The biochemical activity of NC is critical for all of its functions. Several studies have been conducted to gain knowledge in this area. Work done in this thesis sheds light onto two very important aspects of NC's nucleic acid chaperone activity; namely helix-destabilization/unwinding and aggregation/annealing. The former is essentially due to NC's ability to weaken Watson-Crick base-pairing in helical structures while the latter effectively concentrates nucleic acids by causing them to aggregate. Previous work in this lab and others demonstrated that the N-terminal zinc finger (finger 1) of NC is pivotal for helix-destabilizing while the second finger seems to play a lesser role in this activity [91]. The basic approach used in this thesis was to make mutations in finger 1 in an attempt to understand which of the amino acid difference between the two fingers is responsible for their functional differences. A detailed understanding of how NC functions at an amino acid level could potentially help us to uncover ways to target this protein. This could aid in designing molecules that bind tightly to NC and prevent it from interacting with nucleic acids.

Chapter 2 Evaluation of the chaperone activity of HIV-1 nucleocapsid protein at an amino acid level

2.1 Introduction

HIV-1 NC acts as a nucleic acid chaperone. As discussed earlier in Section 1.9, NC's chaperone activity participates in a multitude of functions that it executes throughout viral replication. Upon binding to nucleic acids, it can transiently break base pairs. This enables kinetically trapped nucleic acids to overcome suboptimal conformations. The free bases are now available for re-pairing. NC then catalyzes their rearrangement into optimal conformations with the higher thermodynamic stability. However, the exact mechanism by which this activity is initiated is still not well understood. Such nucleic acid chaperone proteins are also found in prokaryotes and eukaryotes. For e.g., *E.coli* SSB (single-stranded binding protein) has been shown to unwind secondary structures and neutralize negative charges on the phosphate backbone of nucleic acids, thereby catalyzing annealing of complementary DNA strands [92]. Pontius *et al.* have shown that the human A1 hnRNP protein enhances strand exchange and annealing between nucleic acid complements [93]. Williams *et al.* have conducted in depth studies in gp32 (gene 32 protein) found in T4 bacteriophages. Destabilization of duplexes by gp32 and specific binding to single-stranded DNA is vital for phage replication, recombination and repair. However, unlike *E.coli* SSB and T4 gp32 which can bring about complete helix destabilization, HIV-1 NC is relatively a much weaker duplex destabilizer [94]. Thus far, approximately 50 different proteins from virus, eukaryotes, and prokaryotes have been identified as potential nucleic acid chaperone proteins (For a list see: <http://mendel.imp.ac.at/home/Birgit.Eisenhaber/RNA-chaperones/list.html>).

Many attempts have been made to further investigate the role of specific regions in NC responsible for its chaperone activity. Both the flexible backbone and rigid finger regions of NC are of prime importance. However, experiments conducted with finger deletion mutants demonstrate that zinc finger architecture does not appear to be significant for enhancing tRNA annealing and genomic RNA dimerization. *In vitro* tRNA annealing experiments were conducted using NC mutants where the three cysteines residues were replaced by serines (SSHS NC). Results showed that SSHS NC was comparable to wt NC [47]. However, recent studies indicate that this may result from the mutual cancellation of two factors: SSHS NC is less effective than wt NC as a duplex destabilizer, but more effective as a duplex nucleating agent due to increased flexibility [95]. *In vitro* assays conducted by Heath *et al.* have shown that NC finger mutants with little apparent unwinding activity can also stimulate the hybridization of non-structured nucleic acid complements, presumably because wt NC's aggregation effect is not lost [91]. In contrast, results also indicate that the fingers are important for viral packaging [96], [97], minus-strand strong-stop transfer and recombination in general [46], [98]. Cellular infections with NC mutants demonstrated the pivotal role of both fingers in dimer formation and general viability [99]. The unwinding ability of NC also requires zinc fingers, in particular finger one [91]. Studies conducted with mutant NC proteins indicate that 1.1 NC (a mutant NC that has two copies of finger one) and SSHS NC (with an inactivated finger two) retain unwinding activity. In contrast, 2.2 NC (a mutant NC that has two copies of finger two) and SSHS NC (with an inactivated finger one) have little unwinding activity [91]. A finger switch mutant termed 2.1 (where positions of the two fingers are interchanged) also retained partial unwinding activity suggesting that both context and amino acid composition play important roles [91]. Further, truncated NC proteins lacking

the first 12 N-terminal non-finger amino acids (NC 12-55) showed no reduction in helix-destabilizing activity. Single molecule stretching experiments conducted with NC mutants have shown that finger one is crucial for helix destabilization and enhancement of helix to coil DNA transitions [100]. Experiments that measured binding of complementary DNA to the highly structured HIV-1 TAR region (an assay that mimics the first strand transfer event with –ssDNA transferring to the 3' R region) revealed that finger one is vital for optimum annealing. Mutations in finger two did not affect NC's annealing ability to a great extent. Guo *et al.* have demonstrated using NC finger mutants that both context and amino acid content of fingers are critical for both minus and plus-strand transfer reactions. Thus, NC's chaperone activity has been extensively studied and well documented [98].

A clearer understanding of NC's chaperone activity in viral replication would require in depth knowledge of NC's nucleic acid binding properties. Within the mature virion, NC is tightly associated with the genomic RNA. Although it binds to both RNA and DNA, it exhibits preferential binding towards single ss RNA over ss DNA. However, binding affinity of NC to both ss and ds DNA has shown to be similar by fluorescence experiments [101].

Further, it exhibits both general and sequence specific nucleic acid binding properties which are complex and ionic-strength dependant. Surface plasmon resonance experiments have shown that NC binds preferentially to TG repeats in DNA. This was observed with UG regions as well when tested with RNA [102], [103], [104]. Further, NC demonstrates preferential binding towards GNG sequences in single-stranded loop regions [77], [44]. Results indicate that NC exhibits non-specific binding towards t-RNA primers. This is in consensus with the fact that NC does not play a role in t-RNA primer selection [105]. The intrinsic fluorescence of NC's Trp-37 residue and CD-spectral studies has given useful

insights into its binding site sizes [90]. Early studies that monitored NC71 binding to poly(A) have shown the apparent binding site size (n_{app}) to be different under varying protein:nt ratios. The n value was found to be 14 and 8 under low and high ratios respectively. However, upon deletion of 14 residues from the COOH terminus to produce the mature NC55, this variant behavior was resolved and a consistent n_{app} value in the range of 6-7 was obtained. The general consensus is that mature NC's binding site is between 5-8 nucleotides [38], [102], [90], [105], [106], [54], [107], [108].

The ability of NC to bind stoichiometrically to varied nucleic acid structures stems from the high flexibility of its backbone. This has been demonstrated through biophysical experiments by several groups [109], [110], [111], [112], [113], [114]. The large number of polyelectrolyte interactions between NC and the phosphate backbone of nucleic acids supports the hypothesis that the binding event mediated by NC is significantly driven by a large efflux of cations (For e.g.: Na^+ ions). This is clearly demonstrated by NC's salt-dependant *duplex stabilizing* effect [102], [115], [105], [108], [116]. NC bears close resemblance to several cationic ligands like polyamines, polyLysine, cobalt hexamine³⁺ and Mg^{2+}/Ca^{2+} in these aspects. These ligands are highly mobile in nature, bind nucleic acids non-specifically and bring about nucleic acid aggregation similar to NC [117], [118], [119].

A wealth of information about nucleic acid binding properties of NC has been obtained by studying binding interactions between NC and the ψ packaging signal located at 5' end of the viral genome. The ψ site is 120 nucleotides long and contains sequences essential for viral packaging during assembly. Four stem loops namely SL1, SL2, SL3 and SL4 have been proposed to be formed by these nucleotides. Several groups have extensively studied the binding of NC to these stem loops. However, there were significant differences in K_d values

determined. This seemed to depend on the type of NC used, buffer conditions and techniques used for analysis. K_d values of 100-320 nM were observed with SL1 and SL4 indicating weaker binding. The highest binding affinity was reported with SL2 and SL3 with K_d values ranging between 20-30 nM [116]. Assays done with DNA analogues of RNA stem loops revealed that NC binds less tightly to the former [115], [120], [121], [122]. Computational studies have also shown that the maximum contribution towards binding energy in both complexes is from electrostatic interactions between NC-stem loop complexes. Further, while Lys26 residue of NC seems to be very significant for electrostatic binding to SL2, basic residues in the N-terminal helix and finger one seem to be important for electrostatic interactions with SL3 [123]. As mentioned earlier in Sec 1-9, in the NC-SL3 complex, the lysine (Lys3) and arginine (Arg10) residues at the N-terminal zinc finger form a 3_{10} helix. This helix is well packed against the first zinc finger and hence is able to penetrate into the major groove of SL3. However, interactions between SL2 and the 3_{10} helix occur at an A-U-A base triple region in SL2's minor groove [77], [44], [124]. There are lots of similarities between NC interactions with SL2 and SL3. In both cases, intramolecular salt bridges are formed by basic residues which minimize electrostatic repulsions. Further, NC's zinc fingers show preferential binding towards ss loop regions, while the highly positively charged N-terminus binds to the ds stem of the RNA hairpin [77], [44]. Steady-state and time resolved fluorescence studies have also shown that the interaction between Trp37 and guanosine residues depends on its location in the C-terminal zinc finger and correct folding of the finger. While Phe16 contributed vastly to the binding energy, Trp61's (in NC71) contribution was minimal. Further, electrostatic interactions between the N and C termini of NC and SL3 were stabilized by the stem loop structure of SL3 (See Fig. 1-8). Also, Vuilleumier *et al.*

point out that strong binding affinities are not observed during selective binding by NC since it probably recognizes nucleic acids via a small number of sites [115]. Lastly, while nitrocellulose filter binding assays show that NC binds tightly to SL4, gel assays and NMR studies demonstrate otherwise [125], [126].

As described earlier in Sec 1-8, minus-strand transfer generates a minus DNA copy of the RNA genome. Earlier, minus strand transfer was thought to be an intermolecular reaction (occurs on the other genomic RNA) [127]. However, subsequently it was also shown to be intramolecular (transfer occurs on the same genomic RNA template) [128], [129]. As described earlier in Sec 1-9, several groups have demonstrated using *in vitro* assays that HIV-1 NC enhances the rate of minus-strand several fold. HIV-1 NC has been shown to enhance minus strand transfer by facilitating annealing of the R regions and accelerates this step by nearly 3000 fold [72]. The R region contains the highly structured TAR element. Hence, this reaction is most likely brought about by the helix-destabilizing activity of NC [46], [98], [72], [87], [100], [130], [91]. NC mediated unwinding of secondary structures in the R region is thought to be the rate limiting step in this reaction [72]. NC also plays another important role during minus-strand transfer. It inhibits self-priming at the 3' end of –sssDNA. This is a dead-end, non-specific reaction resulting in the formation of TAR-induced fold-back structures that competes with the minus-strand transfer step. Several investigators have conducted in-depth studies in this area. There are conflicting reports on the effect of NC on inhibition of self-priming in the absence of acceptor RNA. However, there is a general consensus that self-priming is greatly reduced when both acceptor RNA and NC are present, resulting in increased rate of strand transfer [131], [68], [46], [98], [132]. This demonstrates a typical case of NC's chaperone activity. Here, the hybrid between –sssDNA and the acceptor

RNA is more stable than self-primed products or either reactant [133]. Biophysical experiments have given useful insights into NC mediated inhibition of self-priming. FRET assays showed that in the presence of acceptor RNA, most TAR DNA hairpins were unfolded. While, in the absence of complementary RNA, although NC enhanced fraying of TAR DNA ends is observed, it was unable to fully unwind the hairpins. This clearly explains why NC is unable to inhibit self-priming under such conditions [134].

Annealing of complementary PBS segments at the 3'ends of minus DNA and +sssDNA constitutes the second strand transfer step. NC plays two important roles during plus-strand transfer. It facilitates removal of the t-RNA primer from minus DNA and stimulates annealing between minus DNA and +sssDNA. NC enhanced plus-strand transfer is another clear manifestation of its nucleic acid chaperone activity. A parallel situation is seen during NC catalyzed minus-strand transfer when it mediates removal of 5'donor RNA fragments from -sssDNA, thereby favoring formation of a more stable RNA-DNA hybrid [82].

However, in addition to these two terminal/end strand transfer events, the viral genome also undergoes transfers internally along the entire genome. They have been shown to occur at a greater frequency during minus DNA synthesis. Internal strand transfers during minus strand synthesis are believed to occur due to several reasons include a break/damage in the genome and secondary structures/specific sequences that hamper the progress of RT along the genome. Such regions are termed pause sites. Refer to Section 1.8 of this thesis for a detailed discussion on internal strand transfer.

Previous results from our lab and others (as discussed above) have shown that the first zinc finger of NC is primarily responsible for its helix destabilizing activity whereas the second finger plays an accessory role by enhancing annealing further [91]. It is not clear what

causes this discrepancy between the activities of the two fingers. There are five amino acid differences between the two zinc fingers of NC in HIV-1 clone pNL4-3. These include (finger one to finger two): phenylalanine to tryptophan (F to W), asparagine to lysine (N to K), isoleucine to glutamine (I to Q), alanine to methionine (A to M), and asparagine to aspartic acid (N to D) at positions 16, 17, 24, 25, and 27 of finger one, respectively (See Fig. 1-7). We wished to determine at an amino acid level the reason for the apparent distinction between the two fingers. Work presented in this thesis evaluates the helix destabilizing activity of NC by analyzing NC mutants in unwinding and annealing assays with structured and unstructured substrates. Nitrocellulose filter binding assays were performed to see if the mutations affect binding to nucleic acids as this may play a role in unwinding activity. Further, this would also give us additional insight into binding affinities of amino acid residues vital for NC's helix-destabilization function. We have also developed *in vitro* systems using purified proteins and virus-derived nucleic acids substrates to model and test internal strand transfer events. Previous results have shown that NC activity is more important for allowing transfer on structured substrates such as the *gag-pol* frameshift region whereas a smaller effect was seen with sequences from relatively unstructured regions. NC finger mutant 1.1 was able to greatly enhance transfer and 2.2 had a lesser effect on *gag-pol*, while both had a similar effect on an unstructured region from the *env* gene [135]. All of the NC mutants were assayed to determine if and to what extent (relative to wt NC) they can stimulate strand transfer. This would further allow us to assess the importance of specific NC amino acids. In this case their direct role in the recombination process, rather than just unwinding will be tested. We hypothesized that there will be a correlation between NC's unwinding activity and stimulation of strand transfer on the *gag-pol* region. Results show that

isoleucine at position 24 and asparagine at position 27 contribute most significantly to the difference between the two fingers with respect to chaperone activity.

2.2 Materials

DNA oligonucleotides for the fluorescence resonance energy-transfer (FRET) assays and hybrid gel-shift annealing assays were purchased from Integrated DNA Technologies (Coralville, IA). T4 polynucleotide kinase, Klenow polymerase and RNase-free DNase I were from New England Biolabs (Ipswich, MA). SP6 polymerase, dNTPs, DNase-free RNase and calf intestinal alkaline phosphatase (CIP) were from Roche Applied Science (Indianapolis, IN). RNase inhibitor (RNasin) was from Promega Corp. (Madison WI). Radiolabeled [γ -³²P] ATP was obtained from Amersham Biosciences (Piscataway, NJ). Sephadex G-25 spin columns were from Harvard Apparatus (Holliston, MA). DE-81 filters and cellulose nitrate membrane filters were from Whatman Inc. (Florham Park, NJ). PCR primers for the strand transfer assay were purchased from Integrated DNA Technologies (Coralville, IA). Plasmid pNL4-3 which is a complete HIV-1 proviral clone derived from NY5 and LAV strains was purchased from NIH AIDS Research and Reference Reagent Program. All other chemicals were from Sigma Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA). Recombinant HIV-RT was purchased from Worthington Biochemical Corp. (Lakewood, NJ). Aliquots of HIV-RT were stored at -80°C , and a fresh aliquot was used for each experiment. All other chemicals were from Sigma Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA).

Preparation of wt and mutant HIV-1 NC proteins: Wild-type HIV-1 NC from either the MN strain (GenBank accession number: M17449), the ARV strain or pNL4-3 was used in this study. Wild-type MN NC was expressed and purified as described [136]. The construct that expresses wt ARV NC (GenBank accession number: K02007) was graciously provided by Dr. Charles McHenry (University of Colorado) and this protein was prepared as described

previously [107]. Wild type and mutant NC proteins from the pNL4-3 sequence (GenBank accession number AF324493) were prepared essentially as described [75]. For the mutant NC proteins, the following sequence changes were made in the gene coding for NC (at the nucleotide level of the pNL4-3 sequence): F16W (t1967g/c1968g), N17K (t1971a), I24Q (a1990c/t1991a), I24E (a1990g/t1991a), A25M (g1993a/c1994t/c1995g), N27D (a1999g/t2001c), and mutants with multiple changes were made with combinations of the above listed changes. The three wt NC proteins differ by no more than five amino acids, which are all functionally conserved. NC aliquots were stored at -80°C in 50 mM Tris-HCl (pH 7.5), 10% glycerol, and 5 mM 2-mercaptoethanol.

2.3 Methods

FRET assay to detect DNA/DNA annealing: The 5' ends of DNA oligonucleotides were tagged with a fluorescein-6-carboxamidoethyl (FAM) label. The complementary DNAs were obtained with a 4-[[[(4-dimethylamino) phenyl]-azo] benzenesulfonicamino (DABCYL) moiety at the 3' ends. Annealing experiments were carried out using a Cary Eclipse fluorescence spectrophotometer (Varian, Inc., Palo Alto, CA). The final concentrations of the FAM and DABCYL DNAs were 5 and 10 nM respectively. FAM and DABCYL complements in the presence or absence of 2 μ M wt NC/mutant NC protein, were separately preincubated in 35 μ l of buffer containing 50 mM Tris-HCl (pH 8.0), 1 mM dithiothreitol (DTT), 6 mM $MgCl_2$, 80 mM KCl, and 25 μ M $ZnCl_2$ for 5 min at 30°C. The solutions were then mixed in a quartz cuvette to start the time course reaction. The time course was monitored over 4 min for the unstructured (0.0dna) and the 5.8dna substrates. With the 9.0dna significant annealing could be observed only by 16min. The FAM molecule was excited at 494 nm and fluorescence emissions were observed at 520 nm. Recordings were made every 10s for the 4 min and every minute for the 16 min time course. The intensity ratio (I_r) was obtained by dividing the peak intensity at every time point (I_t) by the peak intensity observed at time zero (I_0). Plots of I_r vs. time were then constructed. Annealing assays with each wt NC and NC mutants were performed at least three times and an average of the results was used for constructing the plots.

Oligonucleotides used for hybrid gel-shift annealing assay: Some of the NC mutants were also tested in a 7.5rna/dna hybrid system. The RNA for this assay was transcribed from its DNA oligonucleotide pair. One of the DNAs contained an SP6

promoter at its 5'end, followed by the DNA sequence of the 7.5rna. This was named 61mer DNA. The shorter DNA complement of the above without any promoter sequence was called 42mer DNA.

Preparation of 7.5rna/dna hybrid: A hybrid was made between 20 pmoles of the 61mer and 40 pmoles of the 42mer 7.5dna substrates in 50mM Tris-HCl (pH 8.0), 80 mM KCl and 1 mM DTT. The donor/primer reaction was heated to 65-70°C for 5 min and then slow cooled to room temperature. Klenow polymerase was used to fill in the 5'overhang (due to the promoter sequence in the 61mer) in the hybrid. The hybrid made from above was incubated with 200 μ M dNTPs and 10 units of Klenow polymerase. It was then extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol.

Preparation of 7.5rna substrate: Double-stranded hybrid DNA obtained as above was used to generate 7.5rna transcripts with SP6 RNA polymerase using the manufacturer's protocol. The following reagents were used at the indicated final concentrations: 40 mM Tris-HCl pH 8.0, 10 mM DTT, 2 mM spermidine, 6 mM MgCl₂, 2 units/ μ L RNasin (RNase inhibitor) and 1 mM dNTPs. The transcription reaction was carried out at 37°C for two hours. The transcription products were digested with 20 units of RNase-free DNase I for 15 min to remove any remaining template DNA. Equal volume of 2X formamide dye (90% formamide, 10 mM EDTA (pH 8.0), 0.1% xylene cyanol and 0.1% bromophenol blue) was then added. The sample was then denatured by heating between 95-100°C for 3-5 min and electrophoresed on an RNase-free 10% denaturing polyacrylamide gel. Excised gel slices containing the RNA were eluted overnight in 550 μ l of RNA elution buffer (80% formamide, 400 mM NaCl, 1 mM

EDTA, 40 mM Tris-HCl (pH 7.0)). The eluate was then filtered through a 0.45 μ m cellulose acetate syringe filter and precipitated with two volumes of chilled absolute ethanol.

Dephosphorylation of 7.5rna: 50 pmoles of RNA was dephosphorylated with CIP for one hour at 37°C and processed according to the manufacturer's protocol. It was then extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. The 7.5rna was then quantified by measuring absorbance at 260 nm in a Pharmacia Biotech Gene Quant II RNA/DNA spectrophotometer (Piscataway, NJ).

5'end-labeling of 7.5rna: Fifty pmoles of the dephosphorylated 7.5rna was labeled at the 5'-end with [γ -³²P] ATP using T4 polynucleotide kinase according to the manufacturer's protocol. The RNA was then gel-purified on a 10% RNase-free denaturing polyacrylamide gel, excised, eluted and precipitated as described previously. It was then resuspended in 70 μ l of 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0) (TE buffer). The RNA was then quantified spectrophotometrically as described earlier.

RNA/DNA hybrid gel shift annealing assay: The end-labeled 42mer 7.5rna and its DNA complement were separately subjected to a heat snap reaction. This involved heating them to 90°C for 3 min and placing them on ice immediately. The substrates were allowed to incubate on ice for at least 5 min. The final concentrations of RNA and DNA were 5nM and 10nM respectively. The 7.5rna and DNA complements were separately preincubated in the presence or absence of 2 μ M wt/mutant NC proteins in 50 mM Tris-HCl (pH 8.0), 1 mM DTT, 0.1 mM EDTA (pH 8.0), 6 mM MgCl₂, 80 mM KCl, and 100 μ M ZnCl₂ for 2 min at 37°C. 17 μ l of DNA/NC solution was added to 90 μ l of RNA/NC solution to start the time course reaction. Fifteen μ l aliquots were taken out at 0, 0.25,

0.5, 1, 2 and 4 min and stopped by the addition of 7.5µl of stop solution (20% glycerol, 0.25% bromophenol blue, 0.2% SDS, 20 mM EDTA (pH 8.0) and 0.4mg/ml yeast tRNA). They were then incubated for an additional one min at 37°C before being transferred to ice.

Gel electrophoresis: The samples were then resolved on 15% native, polyacrylamide gels. Dried gels were imaged and quantified using a Bio Rad GS-525 phosphorimager.

Annealing assay with 9.0dna: Gel-shift annealing assays were also carried out with the 9.0dnas which were used in FRET assays. Only some mutants were analyzed in order to confirm the findings got from FRET assays. In this case, the 9.0dna containing the DABCYL group at the 3' end was 5' end-labeled with [γ -³²P] ATP as described earlier. The final concentrations of the labeled DABCYL DNA and its complement were 5nM and 10nM respectively. The annealing assay was performed as before except that in this case the annealing reaction was carried out over 16min. Fifteen µl aliquots were taken out at 0, 1, 2, 4, 8 and 16 min and stopped as before. Samples were then electrophoresed, dried and imaged.

0.0dna end labeling for filter binding assay: Fifty pmoles of the 0.0dna was labeled at the 5'-end with [γ -³²P] ATP using T4 polynucleotide kinase according to the manufacturer's protocol. The labeled DNA was then passed over a hydrated Sephadex G-25 spin column to remove any random dNTPs, and processed according to the manufacturer's protocol, then stored at -20°C.

End labeling of 16.3dna for filter binding assay: The 16.3dna (9.0dna without the DABCYL group) without the promoter sequence (42 mer) was used in these experiments.

Fifty pmoles of the DNA was end-labeled with [γ - 32 P] ATP as described above and used in filter binding assays.

Nitrocellulose filter binding assay to monitor binding of wt and mutant NC proteins to nucleic acids: Figure 2-17 shows a schematic diagram illustrating the nitrocellulose filter binding assay. Whatman nitrocellulose membrane filters with a pore size of 0.2 μ m were presoaked for 15 min in nitrocellulose binding buffer (NB buffer: 50 mM Tris-HCl (pH 8.0), 1 mM DTT, 6 mM MgCl₂, 80 mM KCl, and 25 μ M ZnCl₂). NC (0.047-2 μ M) was mixed with 0.0dna (1 nM) in 10 μ l of NB buffer plus 0.1 μ g/ μ l BSA and incubated for 5 min at room temperature. The entire reaction was spotted onto the center of a freshly-pres soaked nitrocellulose membrane filter. The filter was then subjected to vacuum and washed 3 \times with 1 ml of wash buffer consisting of 10 mM Tris-HCl (pH 8.0) and 10 mM KCl. The filters were then air dried. The dried filters were counted using a LKB Wallac 1209 Rackbeta liquid scintillation counter. The fraction of the total substrate that bound to the nitrocellulose filter was then calculated as follows: The counts obtained for each concentration of NC used was initially subtracted from background (determined using a reaction without NC). This value was then divided by the total counts added to the reaction (calculated by counting a reaction applied to a DE81 filter that was not presoaked or washed) to obtain the fraction of the total substrate bound. A plot of fraction substrate bound vs. NC concentration in the reaction was then constructed for each NC protein.

PCR amplification of DNA substrates for strand transfer assays: Two sets of primers were designed to generate the donor and acceptor RNAs derived from the pNL4-3 plasmid. An SP6 promoter sequence (shown in bold) was added to the forward primer

along with 5 additional non homologous nucleotides (shown in italics) to prevent transfer of DNA products from the end of the donor to the acceptor (See Table 2-5). One hundred pmoles of each primer was used in the PCR reactions in a volume of 100 µl including 0.1 µg plasmid DNA, 5 units Taq polymerase, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgOAc. The cycling parameters used were as follows: 35 cycles of successive denaturation, annealing and extension reactions were carried out for 1 min each at 94, 50 and 72°C, respectively. This was followed by one 5 min extension cycle at 72°C. The PCR products were then resolved on an 8% native polyacrylamide gel. Excised gel slices containing DNA were eluted overnight in 550 µl of 10 mM Tris-Cl (pH 7.5), 1 mM EDTA. The eluate was filtered through a 0.45 µm cellulose acetate syringe filter and DNA was precipitated with two volumes of ethanol and 1/10th volume 3 M NaOAc (pH 7.0) and resuspended in water.

Preparation of RNA substrates: DNA products obtained by the PCR method described above were used to generate donor and acceptor RNA transcripts with SP6 RNA polymerase using the manufacturer's protocol. Approximately two µg of the purified PCR DNAs were used. The transcription products were digested with 20 units of RNase free DNase I for 15 min to remove any remaining template DNA. Reactions were then extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. The pellet was resuspended in 70 µl of water and centrifuged through two successive hydrated G-25 spin columns and processed according to the manufacturer's protocol. The RNA was then quantified by measuring absorbance at 260 nm in a Pharmacia Biotech Gene Quant II RNA/DNA spectrophotometer (Piscataway, NJ). The integrity of the RNAs was analyzed to ensure it was comprised of predominately fully

extended RNAs by running approximately 40 pmoles on an 8% denaturing polyacrylamide gel and staining with ethidium bromide.

End labeling of PCR primer for strand transfer assay: Fifty pmoles of the DNA primer that bound specifically to the donor RNA was labeled at the 5'-end with [γ - 32 P] ATP using T4 polynucleotide kinase according to the manufacturer's protocol. The labeled primer was then passed over a hydrated Sephadex G-25 spin column to remove any random dNTPs and stored at -20°C.

Preparation of RNA/DNA hybrid: The donor RNA and the labeled primer DNA were mixed in 10 μ l of 50 mM Tris-HCl (pH 8.0), 1 mM DTT and 80 mM KCl. The donor:primer ratio used was 1:5. The donor/primer reaction was heated to 65-70°C for 5 min and then slow cooled to room temperature.

Strand transfer assay and NC time course reaction: Hybrids from above were preincubated with acceptor RNA in the presence/absence of wild-type or mutant NC proteins for 1 min in 42 μ l of reaction buffer (described below). HIV-1 RT (8 μ l at 0.5 pm/ μ l) was added to start the time course. Final concentrations of reagents in the reactions were: 80 nM RT, 2 nM donor-10 nM primer, 10 nM acceptor, 2 μ M NC, 50 mM Tris-HCl (pH 8.0), 80 mM KCl, 6 mM MgCl₂, 5 mM AMP (adenosine monophosphate) (pH 7.0), 1 mM DTT, 25 μ M ZnCl₂, 100 μ M dNTPs, and 0.4 units/ μ l RNasin. Six μ l aliquots were taken out at 2, 4, 8, 16, 32 and 64 min and stopped by the addition of 4 μ l of 25 mM EDTA (pH 8.0) and 5 ng of DNase-free RNase and incubated for additional 15 min at 37°C. Two μ l of proteinase-K solution (2mg/ml proteinase-K, 10 mM Tris-HCl (pH 8.0), 15 mM EDTA (pH 8.0), 1.25% SDS) was then added to all of the samples, which were incubated at 65°C for 45 min to digest protein. Twelve μ l of 2X

formamide dye (90% formamide, 10 mM EDTA (pH 8.0), 0.1% xylene cyanol and 0.1% bromophenol blue) was then added to each sample. The samples were electrophoresed on 8% denaturing polyacrylamide gels. Dried gels were imaged and quantified using a Bio Rad FX Pro Plus molecular imager with Quantity One software (Hercules, CA).

Gel electrophoresis: Native 8% polyacrylamide gels (for purification of PCR products) and denaturing 8% polyacrylamide gels containing 7M urea (for resolving strand transfer products) were prepared as described in the manufacturer's protocol and subjected to electrophoresis.

2.4 Results of Fluorescence resonance energy transfer (FRET) and gel shift assays with wt NC and NC mutants

FRET assay to detect nucleic acid annealing: Fluorescence resonance energy transfer (FRET) was used to detect annealing between non-structured and structured nucleic acid complements. FRET is defined as “the non-radiative transfer of photon energy from an excited fluorophore (the donor) to another fluorophore (the acceptor) when both are located within close proximity”. However, FRET can be carried out only under certain conditions:

1. The two fluorophores must be within 10-100Å.
2. The emission spectrum of the donor must overlap with the excitation spectrum of the acceptor.
3. Both donor and acceptor molecules should be in their excited energy states.
4. The transition dipole orientations of the two fluorophores must be nearly parallel.

The efficiency of FRET is inversely proportional to the sixth power of intermolecular separation thereby making it a very useful tool to detect changes in proximity over distances comparable to macromolecular dimensions. When donor/acceptor pairs are different, FRET can be monitored either by the sensitized appearance of fluorescence of the acceptor or by the quenching of donor fluorescence. When they are the same, FRET can be detected by the depolarization of the resulting fluorescence. FAM (fluorescein-6-carboxamidohexyl) and DABCYL (4-[[[4-dimethylamino) phenyl]-azo] benzenesulfonicamino) were the fluorophore and quencher molecules that we used in our FRET assays. The FAM molecule was excited at 494 nm and fluorescence emissions were observed at 520 nm. The emission spectrum of FAM overlaps with the excitation spectrum of DABCYL (See Fig 2-1). The 5' end of the donor is tagged with FAM and the 3' end of the acceptor

is tagged with DABCYL (See Fig. 2-2). The donor and acceptor were each 42 nucleotides and were completely complementary. As annealing progresses, the fluorescence emitted by FAM would be quenched by DABCYL. Thus, a decrease in fluorescence intensity would correlate to hybridization between the two oligonucleotides.

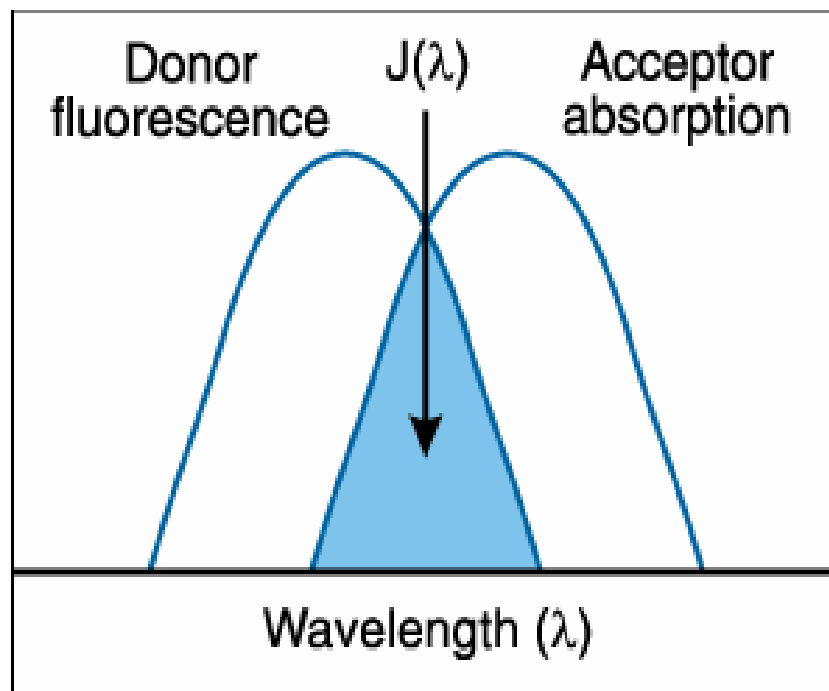


Figure 2-1: Schematic diagram of FRET

Shown is a schematic diagram of FRET. $J(\lambda)$ indicates the spectral overlap integral function. $J(\lambda)$ is plotted as a function over wavelength (λ) in nm. FRET can occur only when the donor emission wavelength spectrum overlaps with the acceptor excitation spectrum. The blue region represents the spectral overlap between the donor fluorescence (emission) and acceptor absorbance (excitation). Figure obtained from “Technical Focus: Fluorescence Resonance Energy Transfer (FRET)” maintained by Invitrogen (<http://probes.invitrogen.com/handbook/boxes/0422.html>).

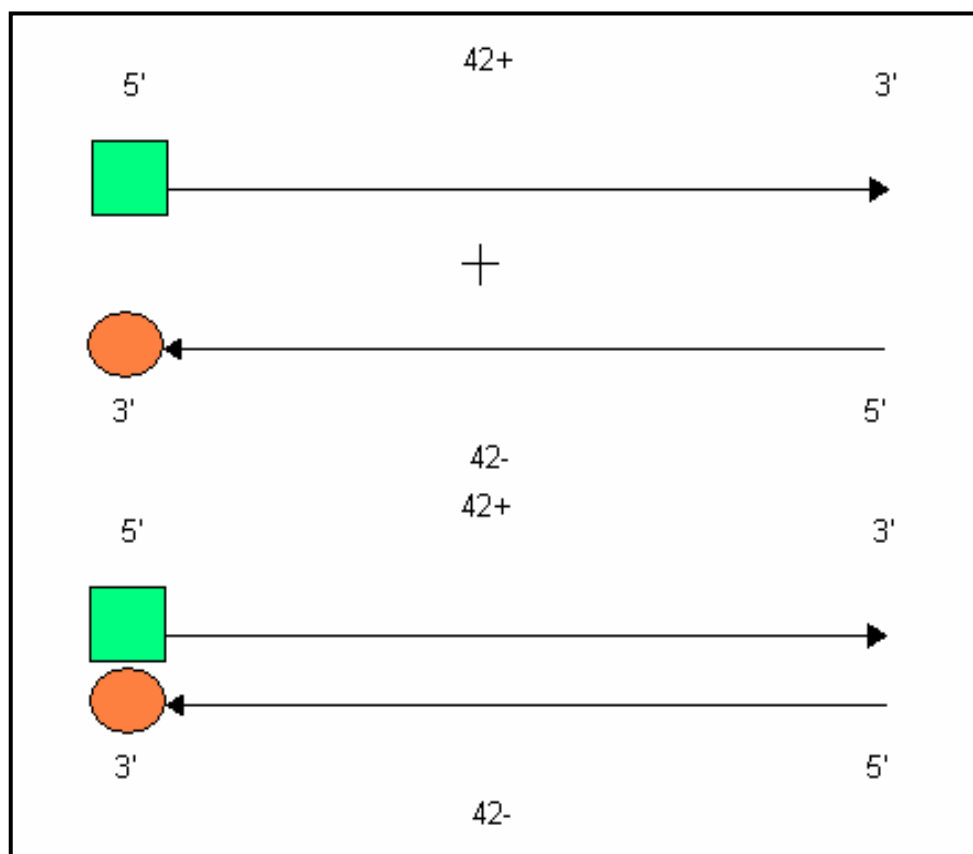


Figure 2-2: Schematic model of FRET assay

Shown is a schematic model of the FRET assay used. The 5' end of the donor oligonucleotide was tagged with a FAM molecule (green square), while the 3' end of the acceptor was tagged with a DABCYL molecule (orange circle). During annealing, the two oligonucleotides come into close proximity with each other. Hence, fluorescence emitted by FAM would be quenched by DABCYL.

Structure of DNA substrates used for annealing assay: All of the DNA substrates were 42 nucleotides in length but differed in sequence composition. These structures were determined using mfold. RNA versions of these substrates had been used in earlier experiments [91]. The predicted folded structures are shown in Fig. 2-3. Note that only one of the two complementary strands is shown for each substrate. The Gibbs free energy of unfolding is indicated next to the corresponding structure and each structure is referred to throughout the text by the energy value followed by “dna” (for example, 9.0dna). All of the substrates formed a stem-loop except 0.0dna which had no predicted structure. The binding strength of the stem in the substrates was increased from one substrate to the next. Successive GT repeats were avoided, since NC has been shown previously to have a preference for GT repeats. Because of the stem-loop structure of the substrate, the complementary strands must be unwound before they can completely hybridize. The assay essentially tests NC’s ability to facilitate annealing by aiding in the unwinding process. Substrate 0.0dna tests the ability of NC to accelerate annealing in the absence of structure. In this case the aggregation/condensation activity of NC is presumably responsible for the observed rate increase.

Annealing assays performed with wt and NC finger mutants: Work previously done in this and other laboratories has shown that the first zinc finger of NC is primarily responsible for unwinding nucleic acid secondary structures (helix-destabilizing activity), while the second finger plays an accessory role [91]. Finger mutants in which the CCHC zinc coordinating amino acids in the fingers were replaced by SSHS were tested for their unwinding activities. Three SSHS mutants were used; SSHS1 had the three cysteines in finger one replaced by three serine residues, SSHS2 had the three cysteines in its second

zinc finger replaced with serine residues, and SSHSd had the cysteines in both fingers replaced with serines. These SSHS finger mutants allowed us to confirm previous findings about roles of the fingers with respect to annealing and helix destabilization. In addition, NC finger mutants 1.1 and 2.2 were also tested (See Sec 2-1). Annealing was detected by FRET as described under Sec 2-3. Assays were performed with mutant and wild-type NC proteins using the various substrates.

The role of the two zinc fingers of NC was previously investigated in our laboratory on an rna/dna hybrid gel shift annealing assay. However we decided to test them using the dna/dna FRET-based system because it is faster and easier to manipulate. The FRET system also allows “real-time” analysis of annealing because the sample is monitored on a scale of seconds rather than minutes in the gel assay. Figure 2-4 and Table 2-1 show the effect of NC finger mutants on 0.0dna. For each NC protein and DNA substrate, a rate constant for complement annealing was calculated by fitting the intensity profile data to a semi-logarithmic plot [72]. This is illustrated for one of the point mutants and wt NC in Fig. 2-11. With the unstructured substrate, the complementary nucleotides annealed very rapidly even in the absence of NC, and NC clearly enhanced annealing even further resulting in about a 2-fold increase in the rate constant (See Table 2-1 and Fig 2-4 also). This demonstrates NC’s ability to enhance annealing even in the absence of secondary structure, presumably by aggregation/condensation. All of the mutants also stimulated annealing with this substrate in comparison to reactions without NC. There were some differences with SSHSd and 2.2 showing the least stimulation; however, all mutants appeared to retain aggregation/condensation activity based on this assay. This is consistent with the idea that this activity results mostly from the highly positively

charged NC backbone amino acids that act to neutralize negative charges on the phosphate backbone of nucleic acids.

Shown in Tables 2-2 and 2-4 are the rate constants for NC proteins on structured DNA substrates (5.8dna and 9.0dna). Very little annealing was observed in reactions without NC and no rate values were obtained. All the finger mutants showed some stimulation. In general the annealing rates were about 4-7 fold slower on 9.0dna than 5.8dna reflecting the greater stability of the former. Consistent with previous results, NC mutants without an active finger one (2.2, SSHS1, and SSHSd) were clearly more defective in helix destabilization. These mutants showed an annealing rate about 20-30% of the wt NC level. These results support earlier findings indicating that the first zinc finger of NC is required for efficient unwinding of strong secondary structures. Further, the various SSHS mutants behave very similarly to the corresponding finger mutants that contain only one of the two zinc finger sequences (1.1 and 2.2).

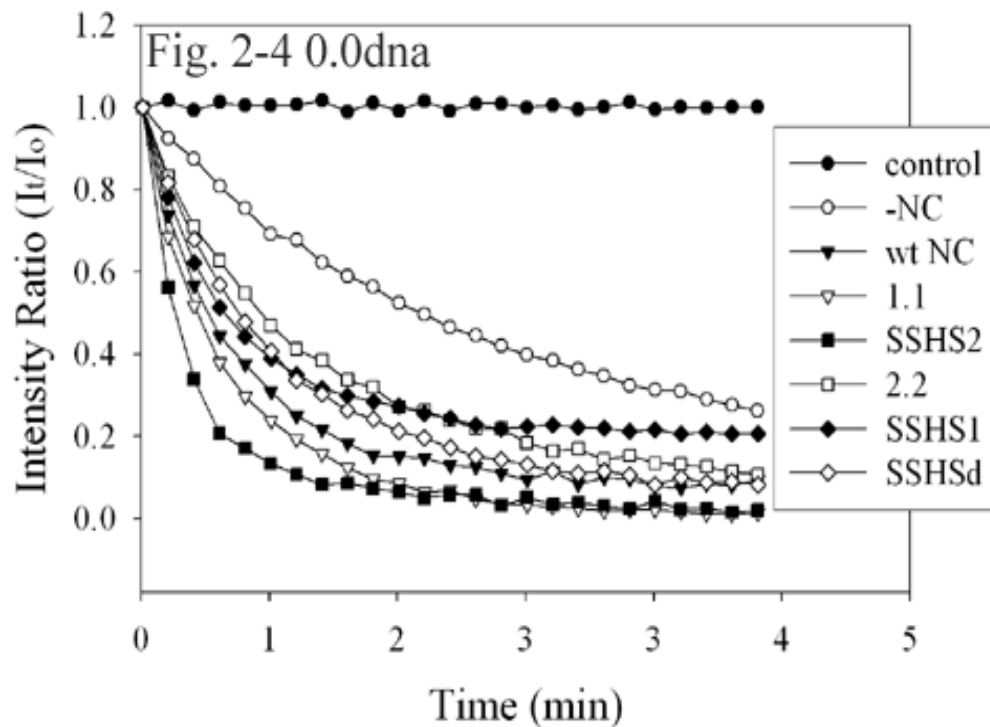


Figure 2-4: FRET assay with NC finger mutants on 0.0dna

FRET assays were carried out at 30°C in the absence or presence of wt or finger mutant NC proteins (1.1, 2.2, SSHS1, SSHS2, and SSHSd) as described in “Methods”. The 5'-FAM derivatized DNA (5 nM final) substrates and 3'- DABCYL derivatized (10 nM final) complements were separately preincubated in the presence of wt or finger mutant NC proteins (2 μ M final concentration). These were mixed to start the annealing reaction. Fluorescence was monitored over time using a fluorescence spectrophotometer. The Intensity ratio (I_r) (fluorescence intensity at each time point (I_t) divided by fluorescence intensity at time zero (I_0)) is plotted versus time (See “Methods”). Recordings were made every 10s throughout the 4 min time course. A control reaction in which no complementary DABCYL DNA was present was also performed. The line shown represents an average from at least three experiments.

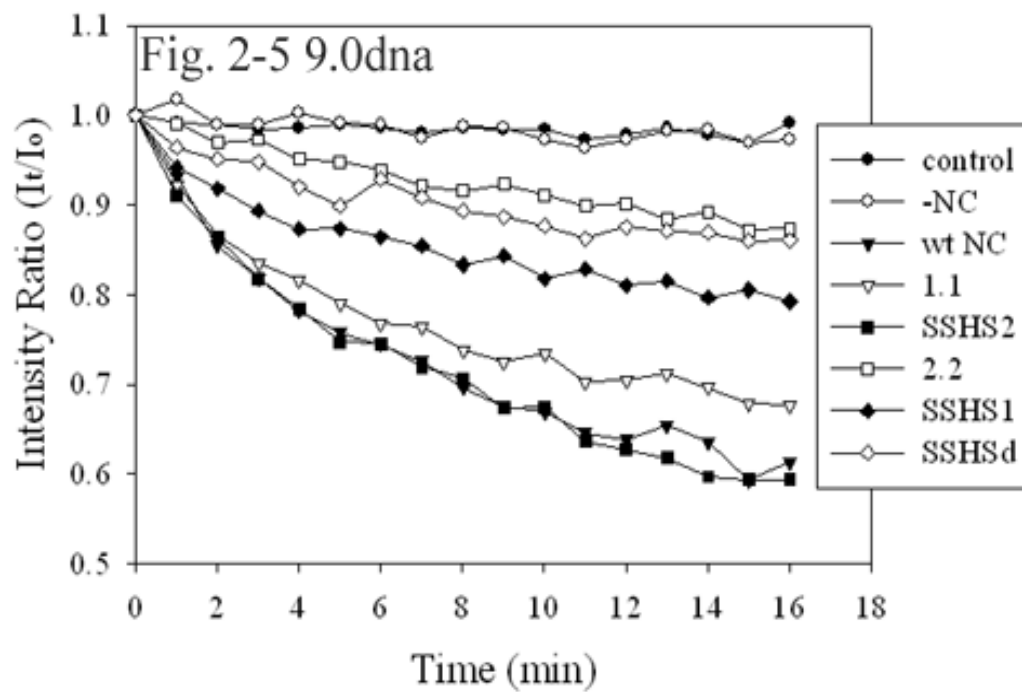


Figure 2-5: FRET assay with NC finger mutants on 9.0dna

Shown above is the FRET assay performed with NC finger mutants on 9.0dna. The assay was carried out as described above (Fig 2-4), except that recordings were made every 1min during the 16min time course.

Table 2-1. Rate constant (<i>k</i>) calculation for NC proteins on 0.0dna substrate		
¹ Name	² t_{1/2}(min)	² k*10⁻³(min⁻¹)
wt NC	0.598	1.16 ± 0.0492
-NC	1.27	0.549 ± 0.0458
SSHS2	0.592	1.18 ± 0.106
SSHS1	0.654	1.06 ± 0.0454
SSHSd	0.813	0.863 ± 0.111
1.1	0.625	1.11 ± 0.0466
2.2	0.634	1.11 ± 0.133
F16W	0.678	1.02 ± 0.0493
N17K	0.586	1.18 ± 0.0193
I24Q	0.602	1.15 ± 0.0321
A25M	0.736	0.957 ± 0.149
N27D	0.760	0.928 ± 0.135
F16WI24Q	0.875	0.792 ± 0.0101
F16WN27D	0.845	0.838 ± 0.171
F16WI24QN27D	0.734	0.949 ± 0.101
<p>1-HIV strain pNL4-3 NC was used as backbone for making NC mutants. The SSHS mutants replace the CCHC zinc binding motif with SSHS in finger 2 (SSHS2), finger 1 (SSHS1), or both (SSHSd). 1.1 and 2.2 are mutants with two copies of finger 1 or 2.</p> <p>2-<i>k</i> (rate constant) values were calculated from the t_{1/2} values by dividing 0.693 by t_{1/2} as described above in Fig. 2-11. Results are an average of 3-4 experiments ± standard deviations (shown for <i>k</i> only). Although experiments were performed over 4 min, only values from the first min were used in calculations due to the rapid kinetics resulting in saturation.</p>		

Table 2-2. Rate constant (*k*) calculation for NC proteins on 5.8dna substrate

¹ Name	² 5.8dna	² 5.8dna (1 mM Mg ²⁺)
-NC	ND	ND
wt NC	0.161 ± 0.0049	0.733 ± 0.019
SSHS2	0.149 ± 0.0231	ND
SSHS1	0.048 ± 0.0009	ND
SSHSd	0.026 ± 0.0024	ND
1.1	0.097 ± 0.0200	ND
2.2	0.038 ± 0.0017	ND
F16W	0.191 ± 0.0287	ND
N17K	0.162 ± 0.0212	0.566 ± 0.049
I24Q	0.069 ± 0.0076	0.281 ± 0.010
A25M	0.109 ± 0.0347	ND
N27D	0.049 ± 0.0017	0.170 ± 0.022
F16WI24Q	0.057 ± 0.0122	ND
F16WN27D	0.058 ± 0.0218	ND
F16WI24QN27D	0.058 ± 0.0355	ND

1-HIV strain pNL4-3 NC was used as backbone for making NC mutants. The SSHS mutants replace the CCHC zinc binding motif with SSHS in finger 2 (SSHS2), finger 1 (SSHS1), or both (SSHSd). 1.1 and 2.2 are mutants with two copies of finger 1 or finger 2, respectively.

2-*k* (rate constant) values were calculated from the *t*_{1/2} values by dividing 0.693 by *t*_{1/2} as described earlier. Results are an average of 3-4 experiments ± standard deviations. Experiments were performed with 2 μM NC, 6 mM MgCl₂ and 80 mM KCl in standard buffer except for the 5.8dna (1 mM Mg²⁺) column. -NC values were only determined for 0.0dna since no significant hybridization was observed for the other substrates in the absence of NC. ND-Not determined.

Table 2-3. Rate constant (<i>k</i>) calculation using varied [NC] on 5.8dna substrate ¹				
NC(μM)	Wild type	N17K	I24Q	N27D
8	0.383 ± 0.206	0.268 ± 0.012	0.178 ± 0.028	0.184 ± 0.097
4	0.335 ± 0.055	0.356 ± 0.025	0.091 ± 0.003	0.078 ± 0.041
2	0.161 ± 0.005	0.162 ± 0.021	0.069 ± 0.008	0.049 ± 0.002
1	0.082 ± 0.027	0.069 ± 0.016	0.032 ± 0.008	0.034 ± 0.007
0.5	0.043 ± 0.001	0.053 ± 0.005	0.034 ± 0.001	0.019 ± 0.011
1- Refer to Table 2-2 for details.				

Table 2-4. Rate constant calculation for NC proteins on 9.0dna substrate		
¹ Name	² t_{1/2}(min)	² k*10⁻³(min⁻¹)
-NC	ND	ND
wt NC	29.2	0.024 ± 0.001
SSHS2	41.4	0.017 ± 0.001
SSHS1	57.6	0.012 ± 0.001
SSHSd	154	0.005 ± 0.001
1.1	37.8	0.018 ± 0.002
2.2	113	0.006 ± 0.001
F16W	6.83	0.102 ± 0.001
N17K	27.5	0.025 ± 0.002
I24Q	56.1	0.012 ± 0.001
A25M	36.0	0.019 ± 0.001
N27D	56.6	0.012 ± 0.001
F16WI24Q	122	0.006 ± 0.001
F16WN27D	122	0.006 ± 0.001
F16WI24QN27D	160	0.005 ± 0.001
<p>1-HIV strain pNL4-3 NC was used as backbone for making NC mutants. The SSHS mutants replace the CCHC zinc binding motif with SSHS in finger 2 (SSHS2), finger 1 (SSHS1), or both (SSHSd). 1.1 and 2.2 are mutants with two copies of finger 1 or finger 2.</p> <p>2-k (rate constant) values were calculated from the t_{1/2} values by dividing 0.693 by t_{1/2} as described earlier. Results are an average of 3-4 experiments ± standard deviations (shown for k only). Experiments were performed with 2 μM NC, 6 mM MgCl₂ and 80 mM KCl in standard buffer. –NC values were only determined for 0.0dna since no significant hybridization was observed for the other substrates in the absence of NC.</p> <p>ND-Not determined.</p>		

Effect of NC point mutants on annealing of structured and unstructured DNAs:

The above experiments confirmed the important role of finger one in helix destabilization (See Sec 2-1). We next wanted to determine what residues in finger one were responsible for its unwinding advantage over finger two. Figure 1-7 shows a schematic diagram of HIV-1 pNL4-3 NC protein. There are five amino acids which differ between the two finger sequences. These include (finger one to finger two): F to W, N to K, I to Q, A to M, and N to D at positions 16, 17, 24, 25, and 27 of finger one, respectively. NC point mutants were constructed where the amino acid residues in finger one were incrementally replaced by those at the corresponding locations in finger two. Figure 2-6 and Table 2-1 show the effects of various NC point mutants on annealing of the unstructured 0.0dna detected by FRET. As with the finger mutants a clear distinction was observed between reactions with and without NC; NC clearly enhances annealing. There was no obvious difference between the point mutants. This was expected given that the more highly mutated finger mutants showed no strong differences with this substrate.

Figure 2-7 and Table 2-2 show the effect observed with the point mutants on 5.8dna. On this more strongly folded substrate clear groupings emerged. Mutants N17K and F16W were able to stimulate annealing as well as wt NC. Mutant A25M showed some reduction in annealing in comparison to N17K and F16W, but was clearly better than N27D and I24Q. The latter two showed about a 60-70% rate reduction in comparison to wt NC.

The same general pattern was observed with the strongest structure, 9.0dna (Fig. 2-8 and Table 2-4). In this case annealing was much slower than with 5.8dna (approximately 7-fold), and the assay was performed over 16 min rather than 4 min as

with 0.0dna and 5.8dna. Once again N17K was similar to wt NC and A25M showed a small reduction in activity. Mutants N27D and I24Q showed reduced stimulation with about 50% the rate of wt NC. Interestingly, the F16W mutant annealed much better than even wt NC showing about a 4-fold increase in annealing rate (Fig. 2-9). This was also the case for an even stronger 42 nucleotide DNA substrate (15.8dna) that folded with a ΔG value of -15.8 kcal/mole (Fig 2-3). On this substrate wt NC stimulated annealing less than half as well as F16W (Fig 2-10).

Optimal magnesium concentrations (6 mM) for reverse transcriptase assays were used in the above described experiments. However, the concentration of non-complexed magnesium in cells may be much lower and results have shown that NC is more active with lower ionic strength. Wild type, N17K, I24Q, and N27D NC proteins were also tested using 1 mM magnesium on 5.8dna (Table 2-2). As expected the annealing rates increased about 3-4 fold with the lower magnesium concentration. Differences between the wt and mutant NC proteins remained consistent at both concentrations, indicating that the magnesium concentration was not a factor in this observation.

An NC titration was also performed using 5.8dna and each of the above mutants (Table 2-3). For wt NC, the rate of annealing was proportional to the [NC] up to 4 μ M NC. Increasing to 8 μ M had little effect. This was also observed with N17K. The more defective mutants (I24Q and N27D) also showed proportional increases but did not saturate at 4 μ M NC. For both mutants rate values at a given NC concentration were always significantly lower than wt NC.

Effect of double and triple mutants on annealing of structured and unstructured DNAs: Results from the single point mutants in annealing assays showed that two of the

mutants, N17K and A25M, were comparable to wt NC with the latter being slightly less stimulatory. In contrast, I24Q and N27D mutants were strongly inhibitory while F16W showed enhanced activity on highly structured substrates. Therefore, of the five different amino acids between fingers one and two, only two (I24 and N27) seemed to be important to finger one's advantage in helix-destabilization. One change (F16W) actually indicated that finger two has a more effective amino acid for helix-destabilization at that position. It is important to note that single point mutations ignore the possibility of "context effects" in contributing to the activities of the fingers. A particular amino acid could function differently in the context of one group of amino acids vs. another. Testing all combinations of mutations in which amino acids in finger one are replaced by corresponding ones from finger two, would require 30 separate mutants, many of which would yield little information. We took a directed approach based on the results from the single mutants. Our working hypothesis was that the changes at positions 24 and 27 cause finger one to lose helix-destabilizing activity and these would have "dominant negative" effects with respect to the apparent gain in activity from the F16W change. This would explain why finger two (as judged from mutants 2.2 and SSHS1) shows little destabilizing activity. To test this, two double mutants (F16W/I24Q and F16W/N27D) containing the positive change along with each negative change, and a triple mutant (F16W/I24Q/N27D) which combined the positive mutant with both negatives were constructed.

On the non-structured 0.0dna substrate (Table 2-1) the double and the triple mutants displayed a very similar annealing pattern with a slightly slower rate than wt NC but a clear increase over reactions without NC. Again, the results suggest that the mutants

retain most of the aggregation/condensation activity of NC. In contrast, the double and triple mutants were all highly defective on the structured substrates (Tables 2-2 and 2-4). These mutants were comparable to 2.2 and SSHSd on the structured substrates. Overall the results support the dominant negative hypothesis and suggest that I24 and N27 are pivotal to the helix-destabilizing function of finger one. Changes in either position to the corresponding amino acid in finger two lead to significant decreases in helix destabilization while changing both essentially mimics replacing finger one with finger two (2.2) or deactivating finger one (SSHS1). These changes also completely mask the strong destabilizing activity of the phenylalanine to tryptophan mutation at position 16 of finger one (F16W).

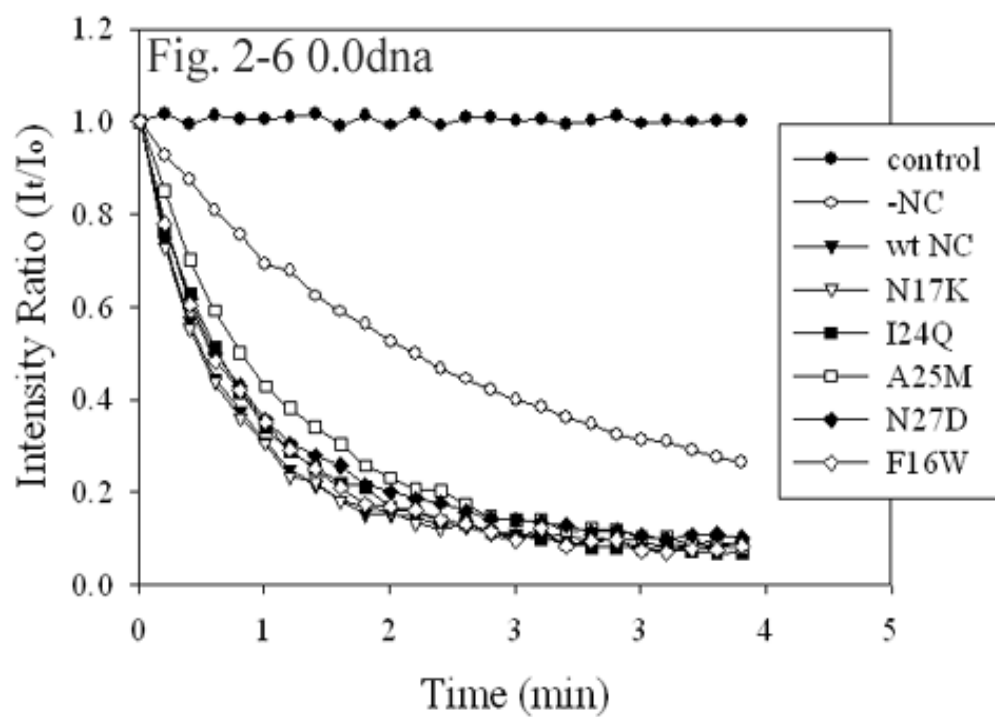


Figure 2-6: Effect of NC point mutants on annealing of 0.0dna

Shown is the FRET assay done with NC point mutants on 0.0dna. Experiments were performed as described earlier (See Fig. 2-4).

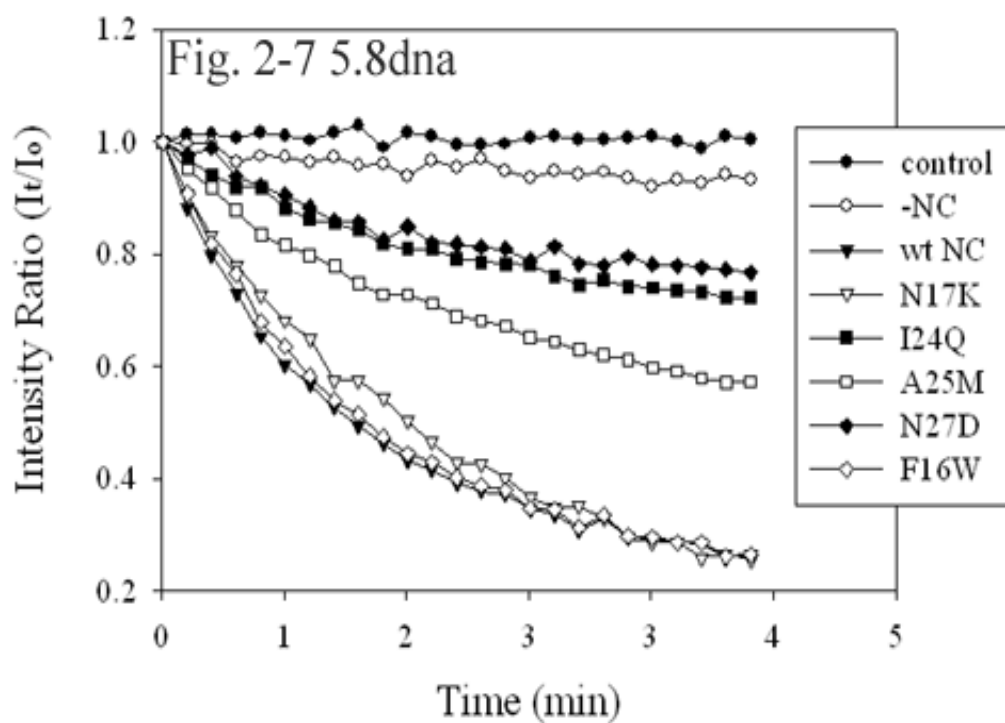


Figure 2-7: Effect of NC point mutants on annealing of 5.8dna

Shown above is the FRET assay done with NC point mutants on 5.8dna. Experiments were carried out as before (See Fig 2-4). Assays with wt NC or without NC are also shown.

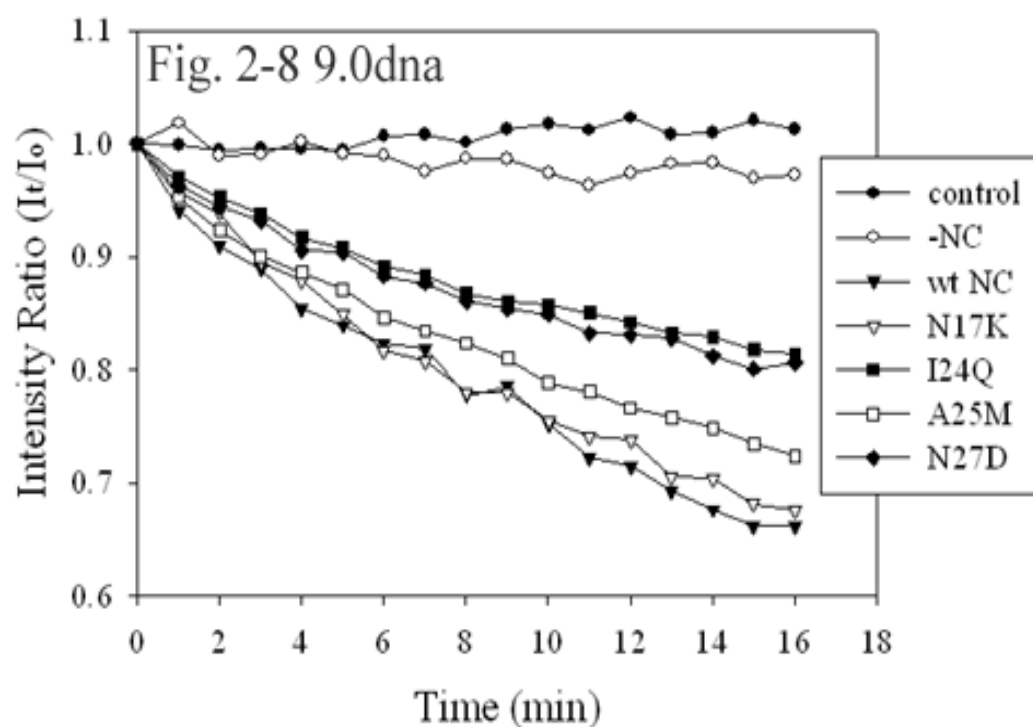


Figure 2-8: FRET assay with NC point mutants on 9.0dna

Shown above is the effect of NC point mutants on annealing of 9.0dna detected by FRET. Recordings were made every min during the entire 16min time course. The assay was performed as described earlier (See Fig 2-4).

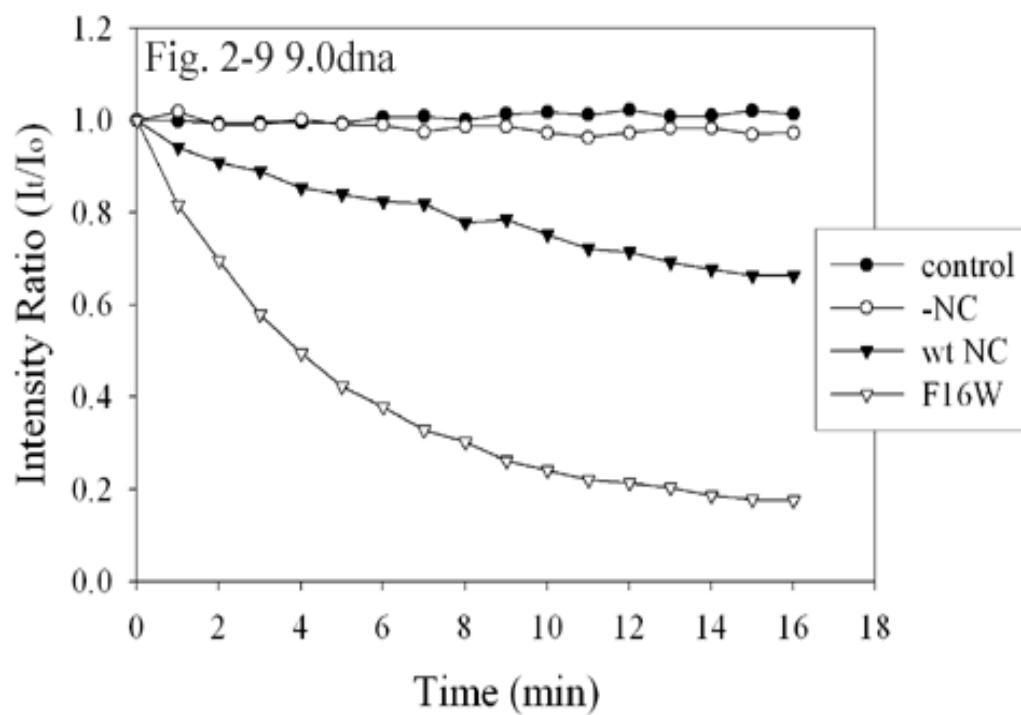


Figure 2-9: Binding of F16W to 9.0dna detected by FRET

Shown above is the effect of the F16W mutant on annealing of 9.0dna detected by FRET. Recordings were made every min during the entire 16min time course. The assay was performed as described earlier (See Fig 2-4).

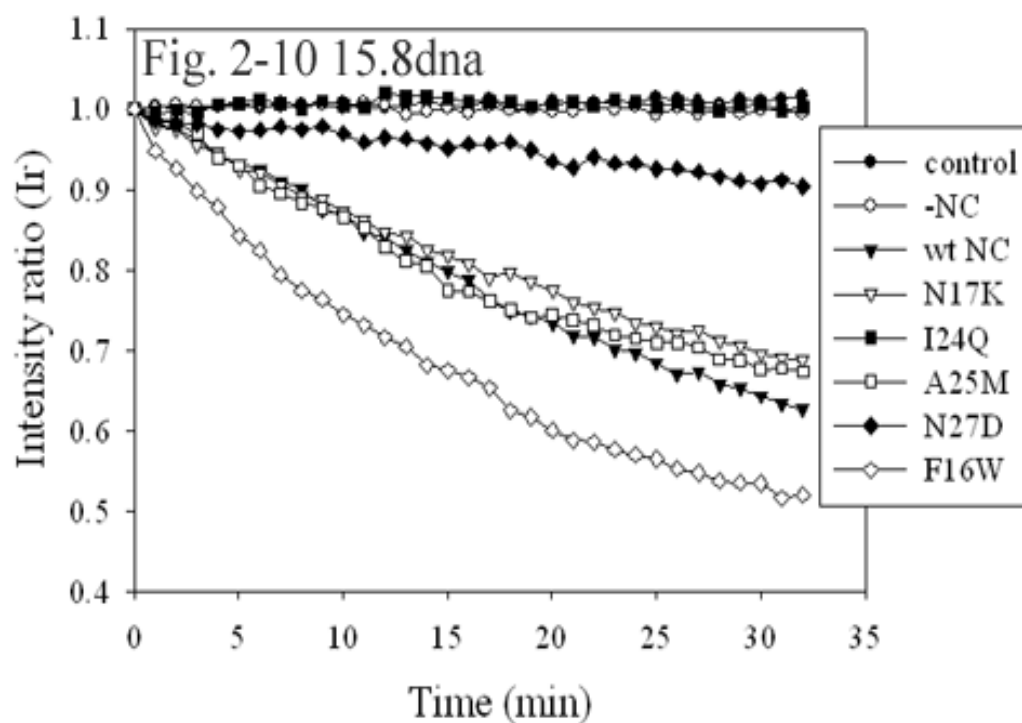


Figure 2-10: Effect of NC point mutants on annealing of 15.8dna

Shown is the effect of NC point mutants on annealing of 15.8dna. The assay was performed as described previously except that recordings were made every min throughout the 32min time course (See Fig 2-4).

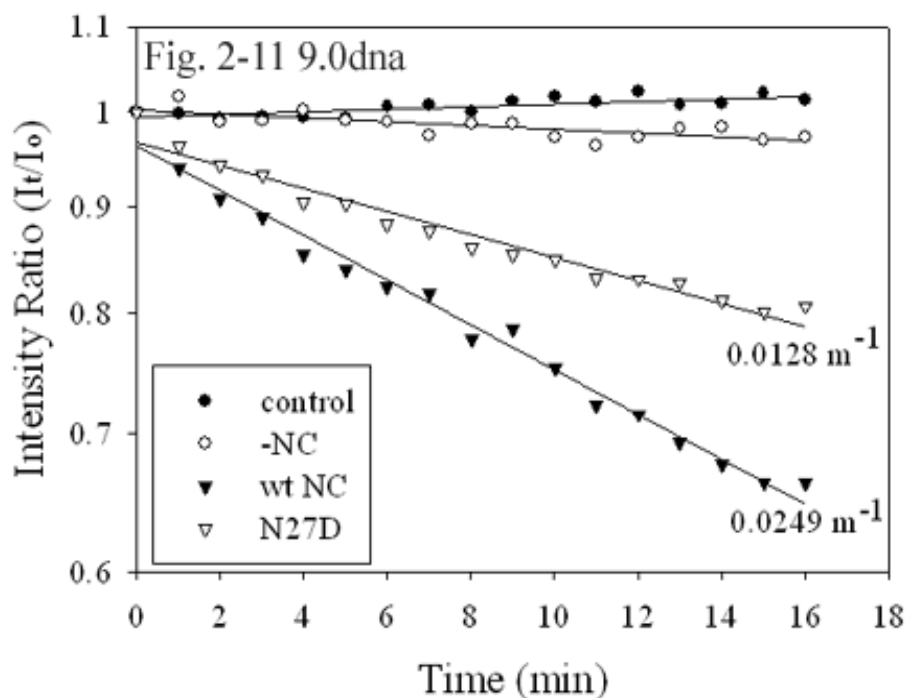


Figure 2-11: Semi-log plot of the intensity profile of N27D mutant

Shown above is a semi-log plot of the reaction with wt and N27D NC constructed from the trace shown in Fig. 2-8. The I_r value from the graph was plotted on a log scale vs. time and the data points were fit to a straight line using linear regression analysis (Sigma Plot). The slope of the line was used to determine the $t_{1/2}$ value of the reaction and the equation $k=0.693/t_{1/2}$ was used to calculate the rate constant k [72]. At least three independent k values were obtained for each mutant on each substrate. Rate constant values for this particular experiment are shown next to each line. The average values \pm standard deviations are shown in Tables 2-1 through 2-4.

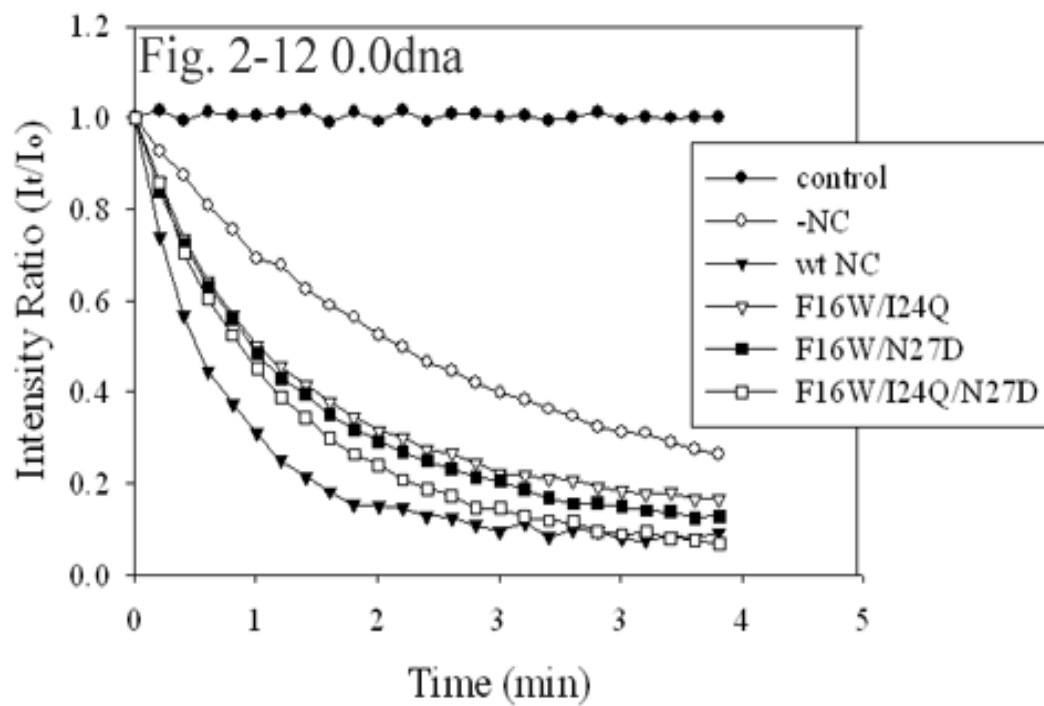


Figure 2-12: Effect of NC double and triple mutants on annealing of 0.0dna

Shown is the FRET assay done with NC double and triple mutants on 0.0dna. Experiments were performed as described earlier (See Fig. 2-4).

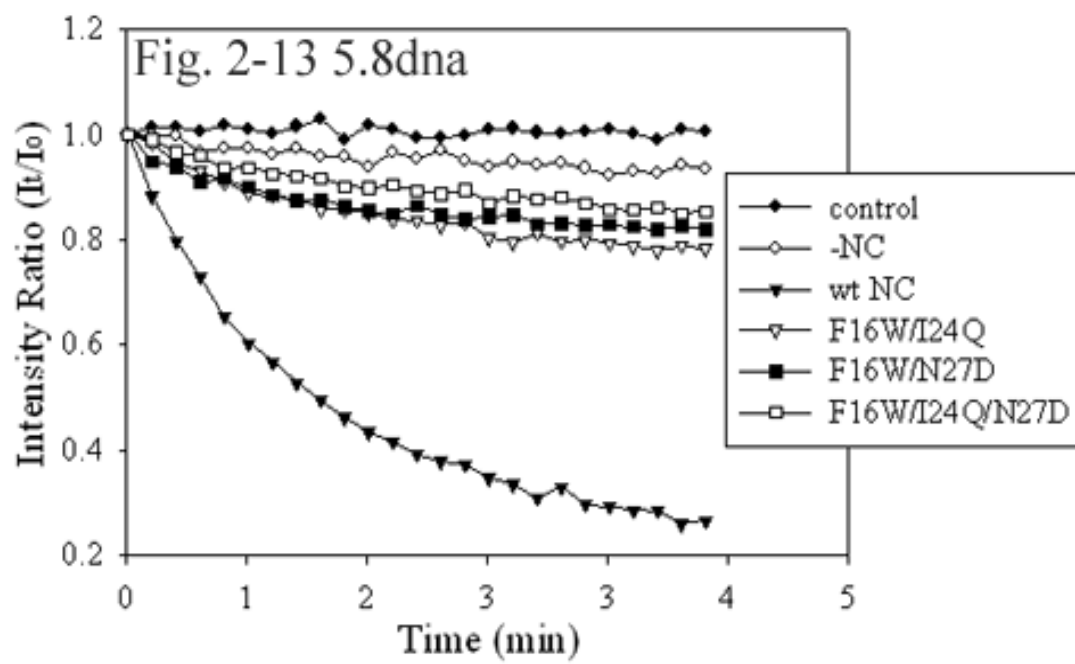


Figure 2-13: Effect of NC double and triple mutants on annealing of 5.8dna

Shown is the FRET assay done with NC double and triple mutants on 5.8dna. Experiments were performed as described earlier (See Fig. 2-4).

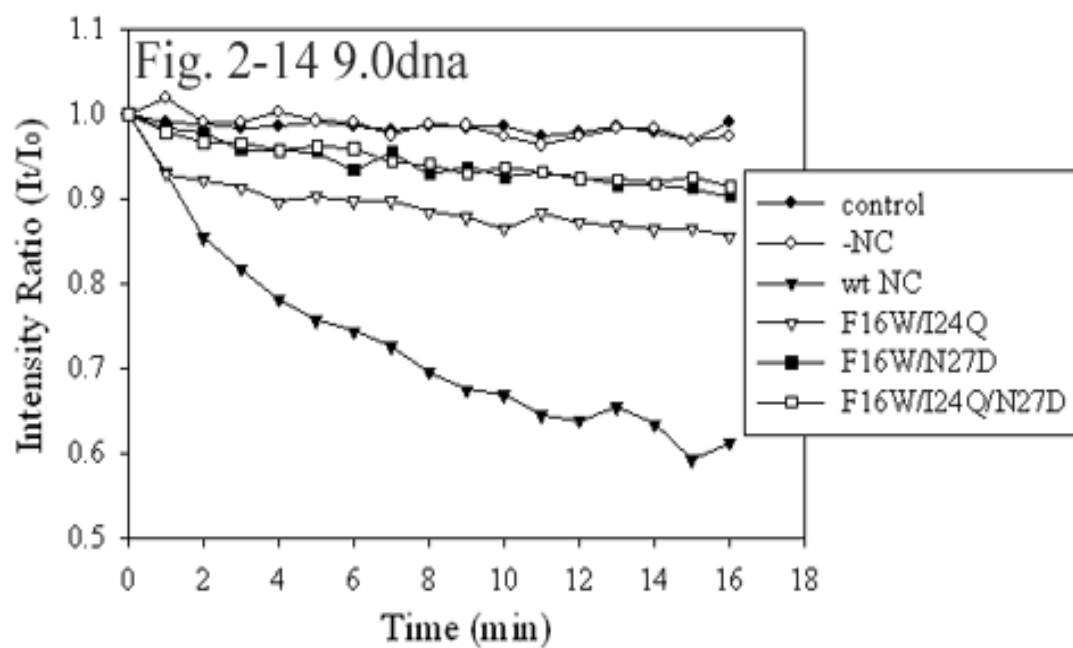


Figure 2-14: Effect of NC double and triple mutants on annealing of 9.0dna

Shown is the FRET assay done with NC double and triple mutants on 9.0dna. Experiments were performed as described earlier (See Fig. 2-4).

Gel-shift annealing assays performed with 7.5rna/dna system: Annealing assays were also completed with some of the NC mutants on the 7.5rna/dna hybrid system as described under “Methods” (See Fig 2-15). This substrate is the RNA form of the 5.8dna substrate prepared as described before. Annealing was monitored between the 7.5rna and its DNA complement over 4 min. Note that this substrate is named 7.5rna due to higher ΔG value of unfolding which reflects more stable folding of RNA vs. DNA. Besides wt NC, the following NC mutants were also tested using the gel-shift assay: SSHS1, SSHS2, N17K, I24E and A25M. N17K, A25M and SSHS2 mutants were able to promote RNA/DNA hybridization close to wt NC levels, while SSHS1 and I24E showed highly reduced stimulation. These observations clearly support our findings from the FRET assays.

Gel-shift annealing assays performed with 9.0dna/dna system: Shown in Fig 2-16 are autoradiograms of annealing assays performed with N17K, A25M, F16WI24QN27D, F16WN27D and F16WI24Q NC mutants on the 9.0dna/dna system. These experiments were carried out as described under the “Methods” section. While N17K and A25M clearly enhanced annealing close to wt NC levels, both double and triple mutants were less stimulatory. Once again, these results correlated with our findings from the FRET assays. The appearance of another band above the DNA/DNA hybrid species can be seen in Fig 2-16. This most likely represents a DNA/DNA dimer since it is more prominent in reactions with one of the oligonucleotide pairs. The moderate stability of dimers, though lesser than hybrids enables them to form quite easily.

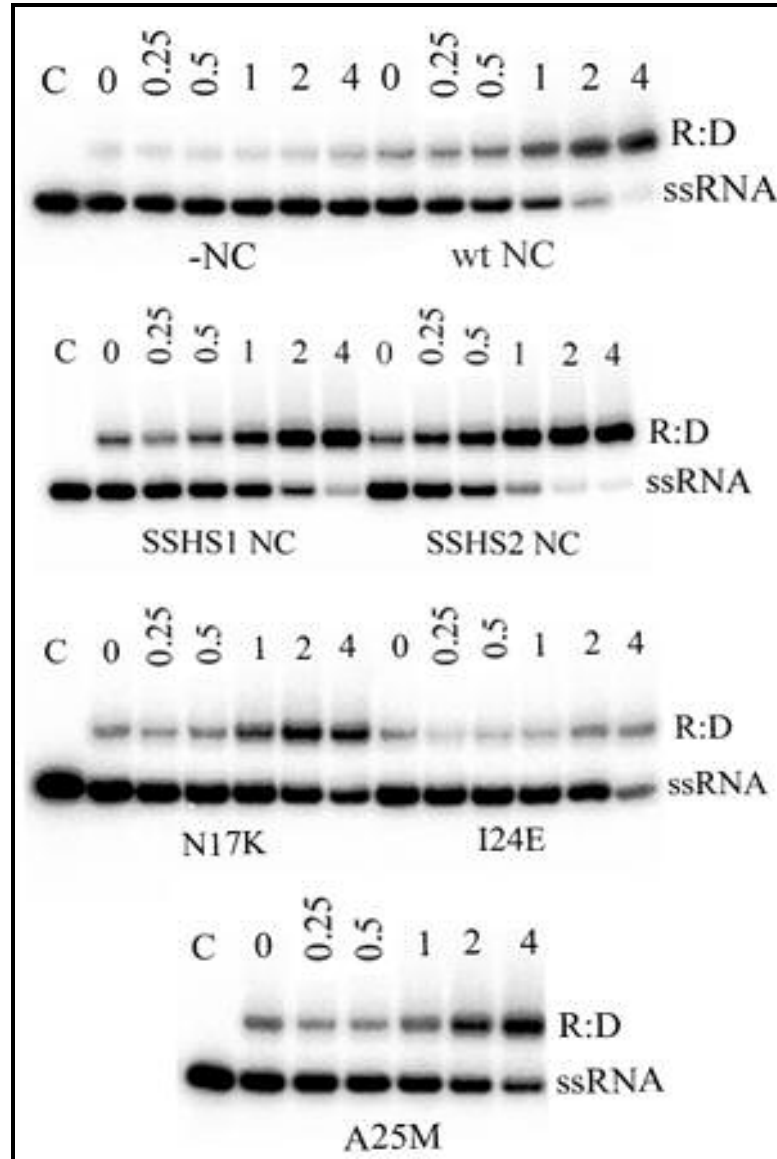


Figure 2-15: Autoradiograms of annealing assays on 7.5rna with NC mutants

Shown in the panel above are autoradiograms obtained from annealing assays performed with NC mutants on the 7.5rna/dna hybrid system. Annealing assays were performed over 4min as described in “Methods”. The time course was monitored at 0, 0.25, 0.5, 1, 2 and 4 min. The positions of ssRNA and RNA/DNA hybrids are also indicated. C denotes control reactions that were carried out in the absence of complementary 7.5dna over the 4min time course.

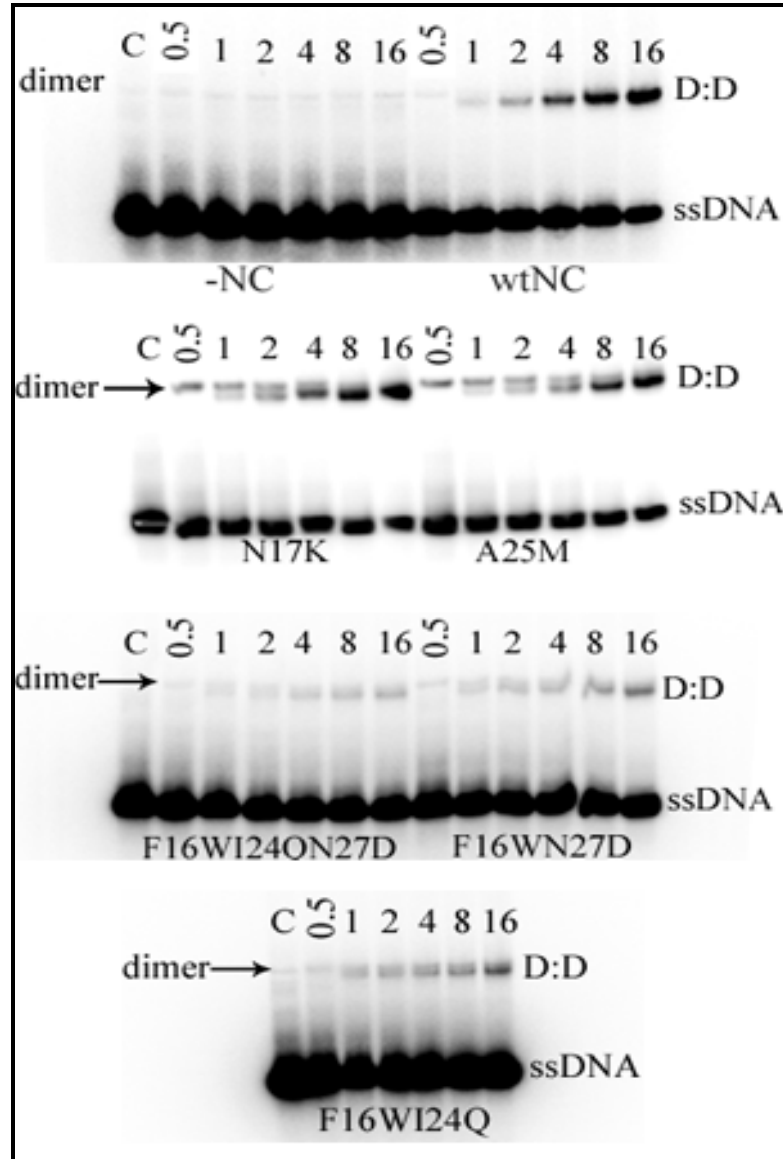


Figure 2-16: Autoradiograms of annealing assays done with 9.0dna

Shown in the panel above are autoradiograms obtained from the 9.0dna/dna hybrid annealing assays. These assays were performed as described before. The positions of single-stranded DNA (ssDNA), DNA/DNA dimers (dimer), and DNA/DNA hybrids (D:D) are also indicated. C is the control sample containing the labeled DABCYL 9.0dna and NC/NC mutant, but no complementary DNA.

2.5 Results of binding affinity analysis for binding of NC and NC mutants to nucleic acid substrates

Binding of NC mutants to 0.0dna detected by nitrocellulose filter binding assay:

One possible reason for the observed defects in some of the mutants could be reduced binding affinity for the substrate nucleic acid. To determine if any of the mutations affected the affinity of NC for DNA, filter binding assays were performed. Graphs obtained by plotting the fraction of total amount of substrate in the assay that bound to nitrocellulose filters at various NC concentrations are shown in Figs. 2-18 through 2-20. The FAM-derivatized complement of 0.0dna made without the FAM group and 5' end labeled with P-32 was used for these assays. SSHSd NC mutant showed the lowest binding affinity (See Fig 2-18). This NC mutant had the cysteine residues in both fingers replaced with serines. This assay allowed us to confirm previous findings about roles of the fingers with respect to DNA binding. Our result clearly supports earlier findings that the presence of zinc fingers is an absolute requirement for nucleic acid binding. The three mutants that are comparable to wt NC in the annealing assays (F16W, N17K and A25M) also show similar levels of DNA binding. In contrast, of the mutants that showed reduced activity in the annealing assays (I24Q, N27D, F16W/I24Q, F16W/N27D, and F16W/I24Q/N27D), only I24Q (Fig. 2-19) was comparable to wt NC in binding. The other four showed significantly lower affinity than wt NC (Fig. 2-20). Consistent with these results, finger mutant 2.2 but not 1.1 showed reduced binding in the assay (data not shown). It was interesting that the double mutant F16W/I24Q showed reduced binding despite neither of the single mutants showing reductions. Overall the results indicate that mutations that reduce NC binding lead to an apparent reduction in helix-destabilizing

activity as expected. However, this activity is not solely determined by binding as I24Q binds similar to wt but is defective in helix-destabilization.

Binding of NC mutants to 16.3dna detected by nitrocellulose filter binding assay:

Mutants N27D and I24Q were also compared to wt NC for binding to 16.3dna in filter binding assays (Figs. 2-21 and 2-22). The apparent binding affinity to this structured substrate was lower for all the NC proteins as more NC was required to retain a comparable amount of substrate on the filters. Consistent with binding to 0.0dna, only N27D showed reduced binding in comparison to wt NC (Fig. 2-21).

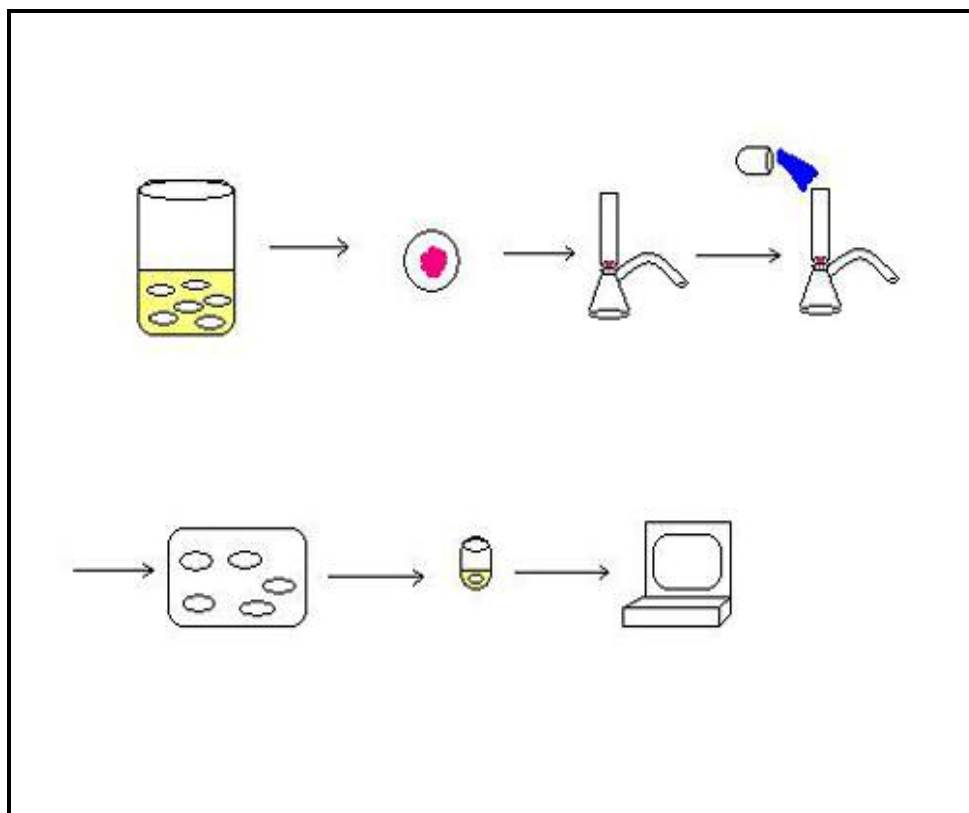


Figure 2-17: Schematic diagram of nitrocellulose filter binding assay

Shown is a schematic diagram of the nitrocellulose filter binding assay. Nitrocellulose membrane filters are first presoaked in NB buffer (50 mM Tris-HCl (pH 8.0), 1 mM DTT, 6 mM MgCl_2 , 80 mM KCl, and 25 μM ZnCl_2) for 15min. NC (0.047-2 μM) is mixed with DNA (1 nM) in 10 μl of NB buffer plus 0.1 $\mu\text{g}/\mu\text{l}$ BSA and preincubated for 5min. The entire reaction is then spotted onto the center of the disc. The filter is then subjected to vacuum and washed 3 \times with 1 ml of wash buffer consisting of 10 mM Tris-HCl (pH 8.0) and 10 mM KCl. The filters are then air dried and counted using a LKB Wallac 1209 Rackbeta liquid scintillation counter.

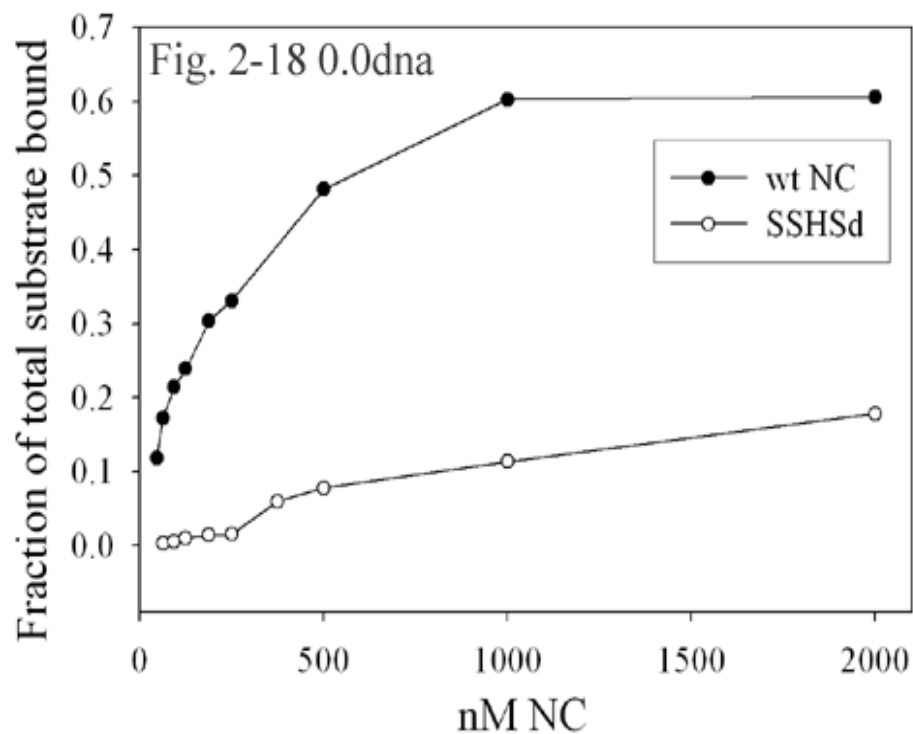


Figure 2-18: Binding of SSHSd to 0.0dna detected by filter binding assay

Shown above is the binding of SSHSd NC mutant to 5'end labeled 0.0dna detected by nitrocellulose filter binding assay. The assay is performed as described previously. The counts obtained for each concentration of NC used was initially subtracted from background (determined using a reaction without NC). This value was then divided by the total counts added to the reaction (calculated by counting a reaction applied to a DE81 filter that was not presoaked or washed) to obtain the fraction of the total substrate bound. A plot of fraction substrate bound vs. NC concentration in the reaction was then constructed for each NC protein. A representative experiment is shown and each assay was repeated at least once.

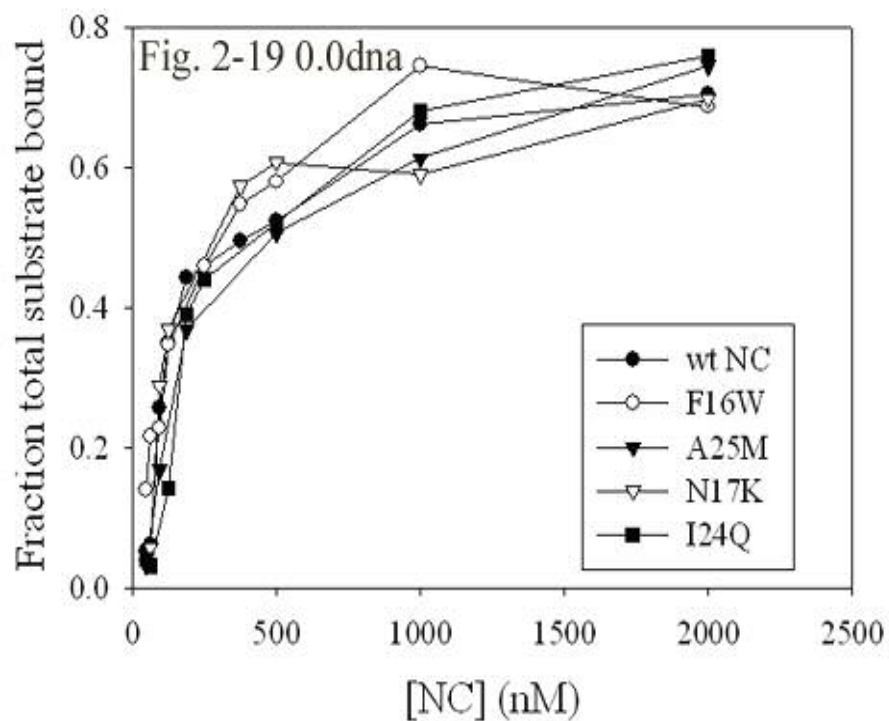


Figure 2-19: Binding of F16W, N17K, I24Q, and A25M NC mutants to 0.0dna

Shown above is the binding of F16W, N17K, I24Q and A25M NC mutants to 0.0dna. The assay is carried out as before and plots are obtained. A representative experiment is shown and each assay was repeated at least once.

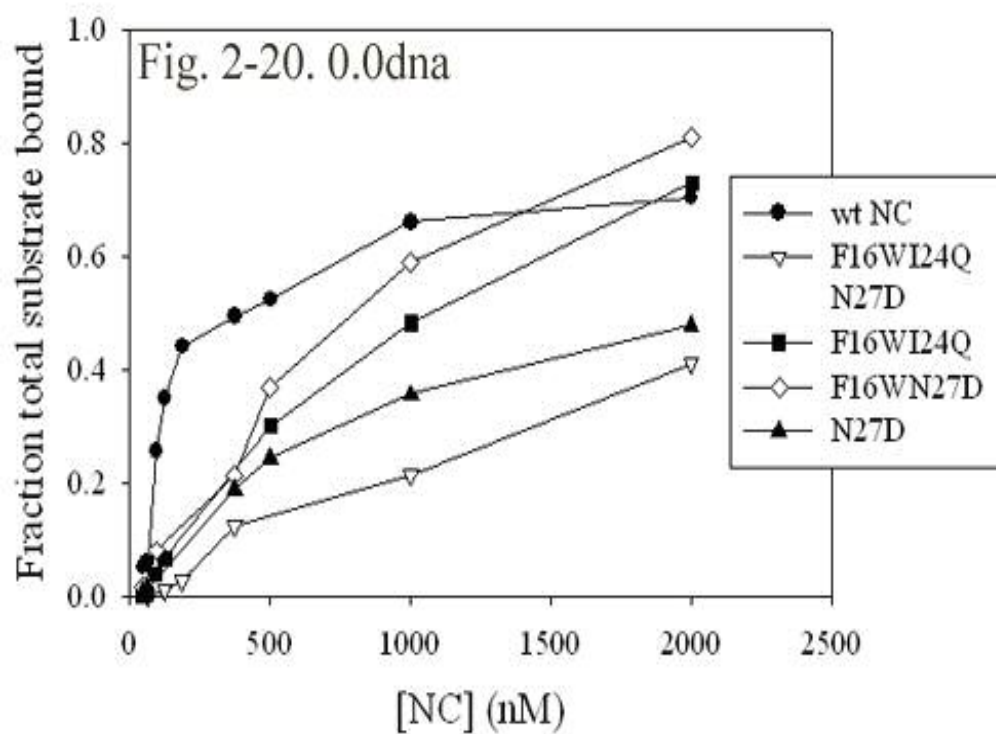


Figure 2-20: Binding of N27D, F16W/I24Q, F16W/N27D, and F16W/I24Q/N27D NC mutants to 0.0dna

Shown above is the binding of N27D, F16W/I24Q, F16W/N27D, and F16W/I24Q/N27D mutants to 0.0dna. The assay is carried out as described previously. A representative experiment is shown and each assay was repeated at least once.

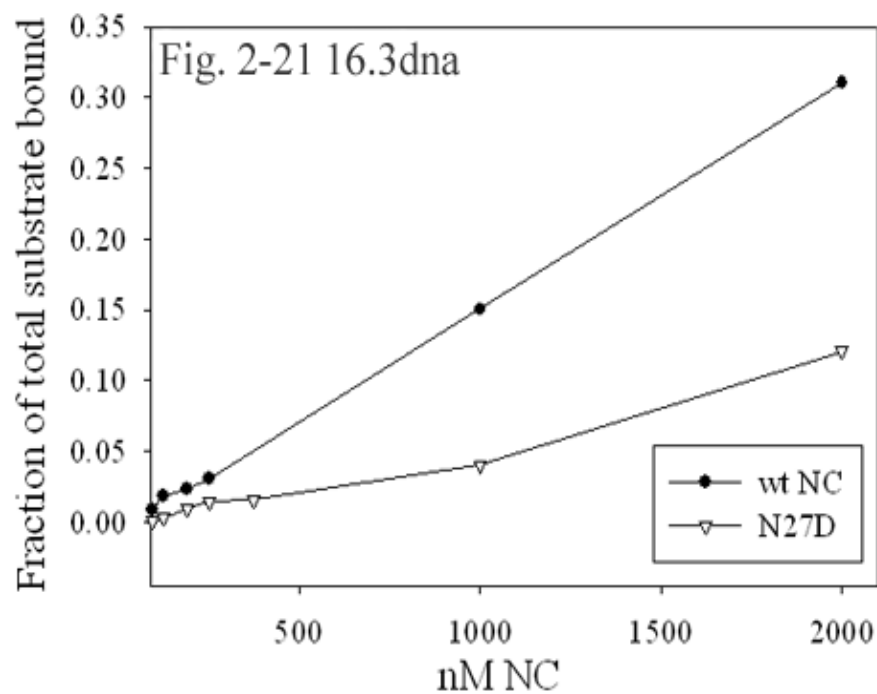


Figure 2-21: Binding of N27D to 16.3dna

Shown is the binding of N27D NC mutant to 16.3dna detected by filter binding assay. The assay is performed as before. The apparent binding affinity to this structured substrate was lower as more NC was required to retain a comparable amount of substrate on the filters. Consistent with binding to 0.0dna, N27D showed reduced binding in comparison to wt NC.

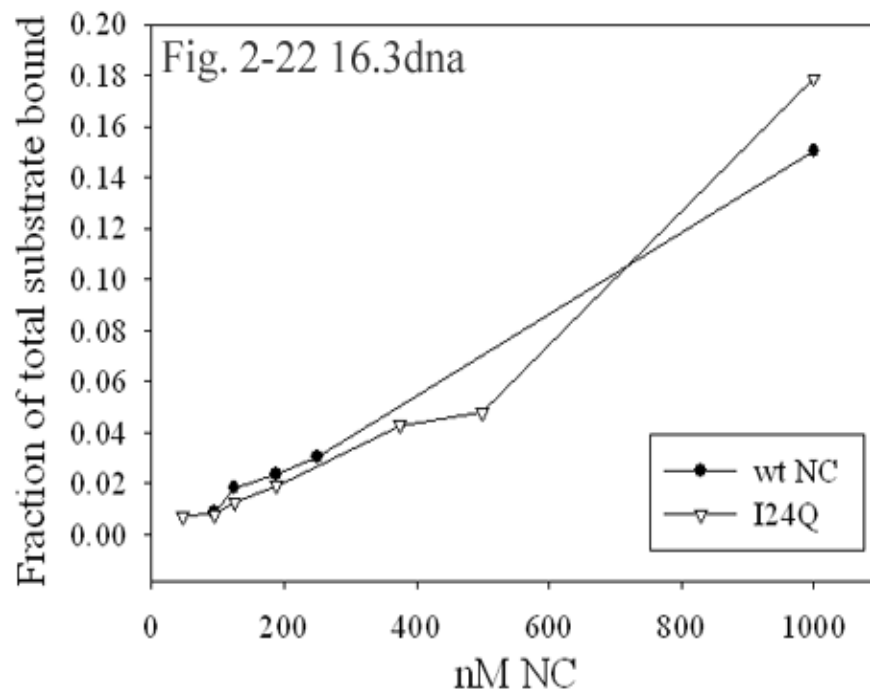


Figure 2-22: Binding of I24Q to 16.3dna

Shown is the binding of I24Q NC mutant to 16.3dna. The assay was performed as described earlier. As with 0.0dna, the binding of I24Q is comparable to wt NC.

2.6 Results of internal strand transfer assays with wt NC and NC mutants

Strand transfer assay: An *in vitro* strand transfer assay has previously been used in our laboratory to investigate recombination in different regions of the genome [135]. Figure 2-23 shows a schematic model of the system. Figure 2-24 depicts a schematic overview of the experimental approaches used. This assay mimics internal strand transfer events that occur during minus DNA synthesis in the viral genome. The donor refers to the template RNA on which DNA synthesis is initiated, while the template on which donor directed DNA products can potentially transfer to is referred to as acceptor RNA. Thus, the donor and acceptor RNA templates represent the two RNA strands of the viral genome. For these experiments, the templates were derived from a highly structured region of the HIV genome that included the *gag-pol* frameshift sequence. DNA synthesis is primed at the 3' end of the donor by a 5' end labeled DNA primer. The donor and acceptor RNA strands are designed to have a 150 base pair region of homology where strand transfer can potentially occur. The transfer zone is indicated by a dotted box. Full length DNA products from DNA synthesis directed by the donor would be 175 nt in length. Products made after transfer and extension on the acceptor would be 197 nt in length. The difference in lengths of the DNA products allowed us to locate their relative positions on denaturing gels. This assay tests for internal strand transfers; hence transfers from the end of the donor were prevented by adding a 5 nucleotide non-homologous region to the 5' end of the donor template. Transfer could also occur between two donor templates. However, this would be negligible since the amount of acceptor RNA used was five times in excess over the donor RNA. Further, fewer donor RNA templates would be available as DNA synthesis proceeds.

The percent of transfer efficiency was calculated by: $\{\text{transfer DNA products (T)} / (\text{transfer DNA products (T)} + \text{full length donor-directed DNA (F)})\} \times 100$. It reflects the amount of DNA primers that get extended to the end of the acceptor versus those that undergo donor extension. This number gives us an actual indication of transfer in terms of total DNA extended, rather than simply giving transfer levels.

Secondary structure prediction using RNAdraw and mfold programs: Secondary structures of RNA substrates were predicted using RNAdraw and mfold programs [135], [137]. Figure 2-25 shows the secondary structure prediction of the donor RNA obtained from the *gag-pol* frameshift region. The acceptor RNA was predicted to have a similar structure as the donor (See Fig 2-26). High ΔG and T_m values indicate presence of strong secondary structures. The *gag-pol* donor RNA was predicted to have a ΔG value of -45.3 kcal/mole and its predicted stem-loops persisted even at temperatures above 55°C. The *gag-pol* acceptor RNA was predicted to have a ΔG value -37.22 kcal/mole. The *A residue* which was determined to be a major pause-site for RT enzyme is also indicated on both folding diagrams.

Effect of NC mutants on strand transfer: Figures 2-27 through 2-30 show autoradiograms obtained from strand transfer assays performed with the *gag-pol* substrate using the different NC mutants. The positions of transfer (T) and donor-directed products (D) are indicated as is the primer position (P). In Fig 2-27, the four sets of assays show reactions (from left to right) without NC, those with wt NC, and those with N17K and A25M NC mutants. The six lanes for each set were reactions stopped after 2, 4, 8, 16, 32, and 64 min from left to right. Graphs of efficiency of strand transfer *versus* time for different NC mutants are shown in Figs. 2-31 through 2-33. Graphs were made from

quantification of experiments as described under “Methods”. All the mutants enhanced transfer considerably compared to reactions without NC. However there was clearly less enhancement, especially at early time points, in comparison to wt NC. Mutants F16W/I24Q, F16W/I24Q/N27D and F16W/N27D showed lesser stimulation than wt NC (Figs. 2-29, 2-30 and 2-31). Mutants N17K, A25M and F16W were comparable to wt NC with N17K showing slightly more stimulation (Figs. 2-27 and 2-32). Mutants I24Q and N27D were less stimulatory than wt NC with N27D showing the lowest stimulation (Figs. 2-28 and 2-33). Again, the differences were more evident at the early time points. The results indicated that even the mutants which show low helix-destabilizing activity enhance strand transfer considerably on this highly structured substrate. However, those with near wt NC levels of unwinding activity showed more stimulation.

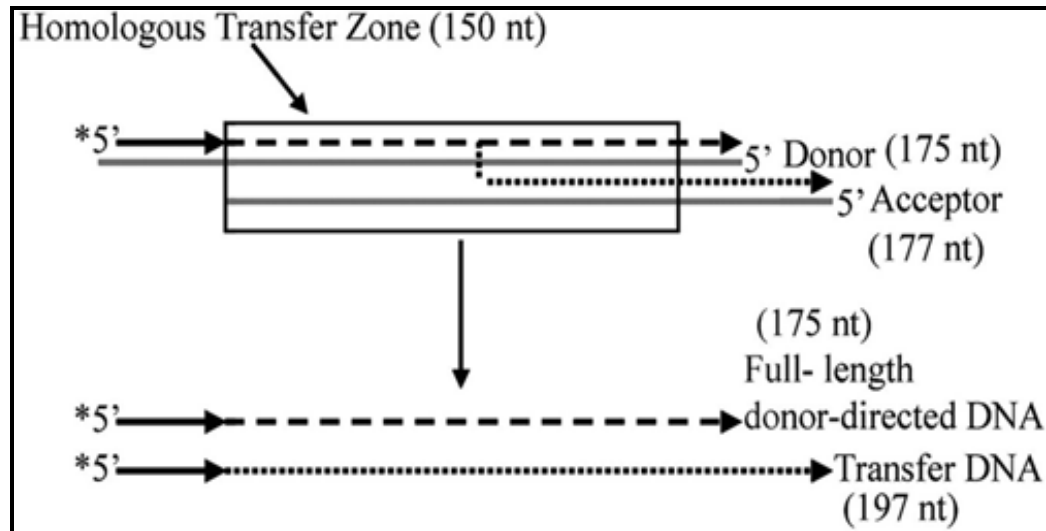


Figure 2-23: Schematic model of strand transfer assay

Shown is a schematic diagram of the strand transfer system. The donor and acceptor RNAs were derived by PCR amplification of the *gag-pol* frameshift region. The 5' end of the 20 nucleotide primer (complementary to the donor) was radio-labeled with P-32 and is indicated with an *asterisk* symbol. The donor and acceptor RNAs were 175 and 177 nucleotides (nt) in length, respectively. They are homologous only in the *boxed region* which spans 150 nt. Full length DNA synthesized on the donor RNA template is indicated by *dotted lines*. Transfer DNA synthesized on the acceptor RNA template is indicated by *broken lines*. Note that the donor-directed DNA product is shorter in length in comparison to the transfer DNA product (Figure obtained from Derebail *et al.* [135]).

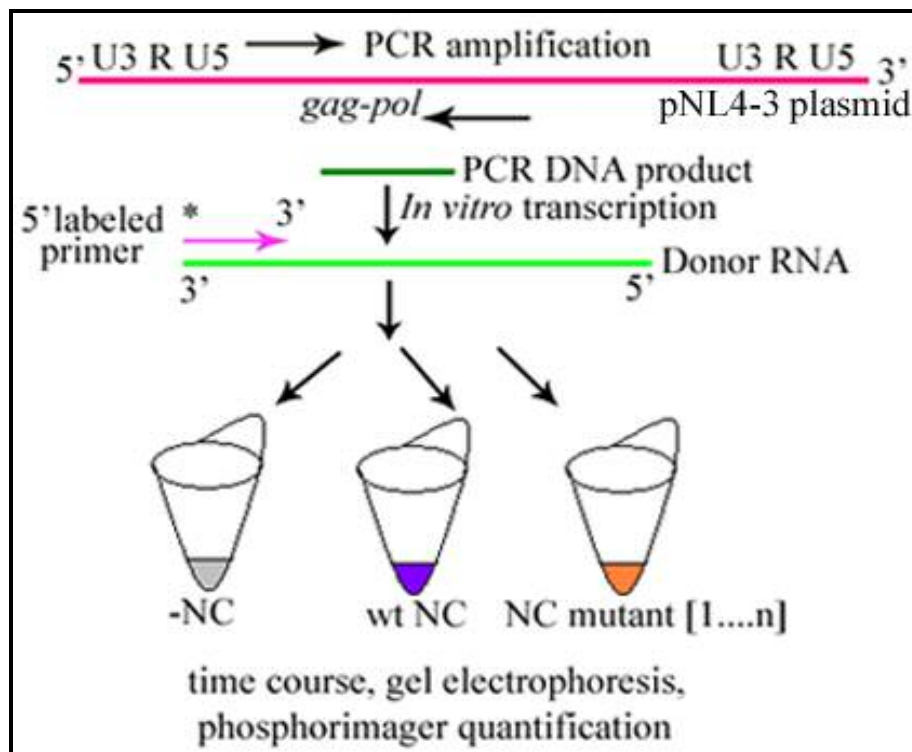


Figure 2-24: Schematic of experimental approaches used

Shown is a schematic diagram illustrating experimental approaches used in the strand transfer assay. The donor and acceptor DNAs (dark green) were derived by PCR amplification using forward and reverse primers (black) from the *gag-pol* frameshift region of the pNL4-3 plasmid (pink). *In vitro* RNA transcription was then conducted using SP6 RNA polymerase to get donor and acceptor RNA templates (only donor RNA is shown in light green). A 5' P-32 end-labeled DNA primer (pink line with *asterix* symbol) synthesized DNA on the donor RNA (primer was complementary to the donor only). The primer/donor hybrid was used in several time course reactions in the presence of acceptor RNA. -NC represents a reaction in which no NC was present. wt NC represents a reaction carried out with wt NC. NC mutant [1....n] represents reactions carried out with all the NC mutants.

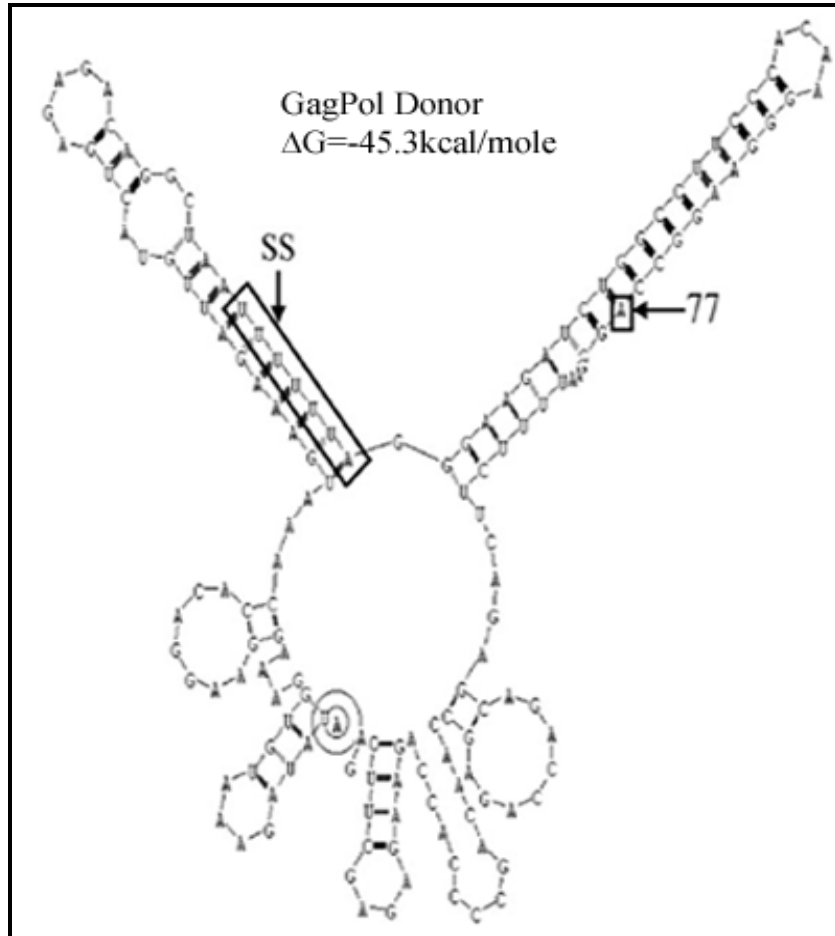


Figure 2-25: Predicted secondary structure of *gag-pol* donor RNA template

RNAdraw and mfold programs were used to predict the secondary structure of *gag-pol* donor RNA template at 37°C and default conditions. Concentric circles indicate the 5' end. The substrate is 155 nt in length (excluding the first 20 nt that bind to the DNA primer). Base-pairs having a greater probability of formation are denoted by thicker lines. The ΔG value of unfolding is also indicated. 77 besides a boxed A residue denotes the major RT pause site. SS within a box denotes the “slippery site” heptamer sequence (Figure obtained from Derebail *et al.* [135]).

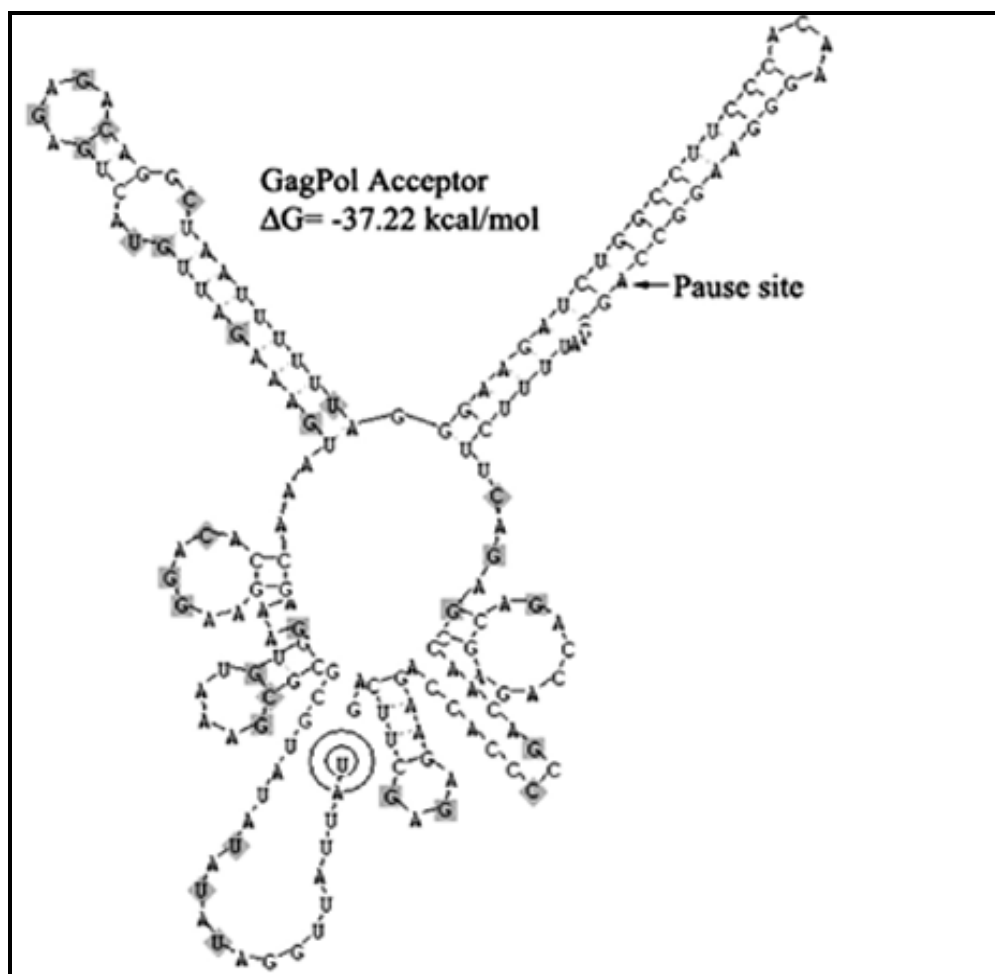


Figure 2-26: Predicted secondary structure of *gag-pol* acceptor RNA template

RNAdraw was used to predict secondary structure of the *gag-pol* acceptor at 37°C and default conditions. As before, concentric circles indicate the 5' end. The ΔG value of unfolding is also indicated. *Boxed G residues* denote T1 RNase cleavage sites. *Diamond A/U residues* denote RNase A cleavage sites. The major pause site for RT (*A residue*) is also denoted (Figure obtained from Derebail *et al.* [137]).

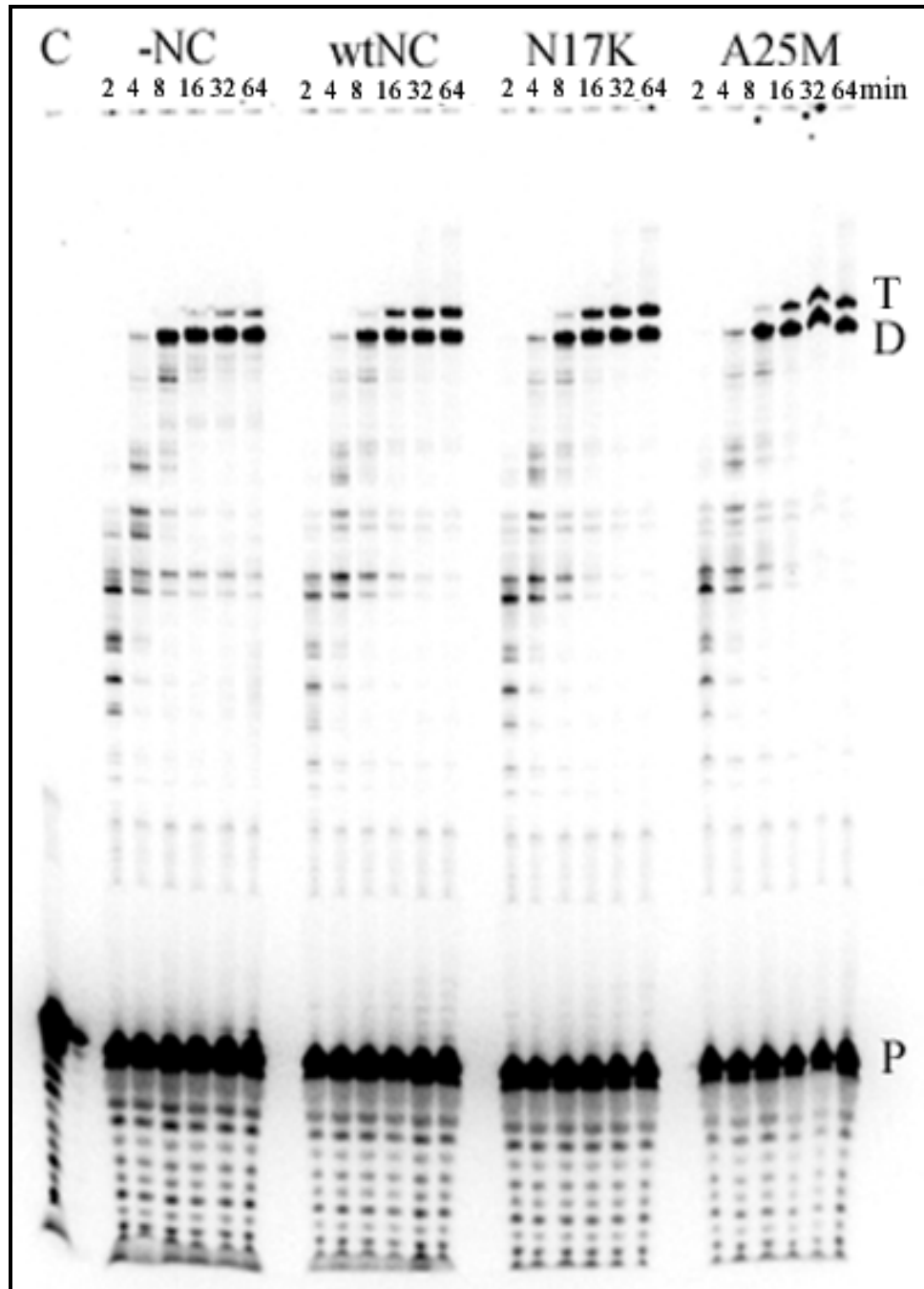


Figure 2-27: Autoradiogram of strand transfer assay with N17K and A25M

An autoradiogram of a strand transfer assay carried out with the N17K and A25M mutants is shown. The assay was performed as described under “Methods”. The four sets of assays shown from left to right were reactions carried out without NC, with wt NC, and those with N17K and A25M mutants. Aliquots from each reaction were stopped at 2, 4, 8, 16, 32, and 64 min and analyzed on an 8% denaturing polyacrylamide gel. Positions of the donor-directed and transfer products and the primer are indicated by D, T, and P respectively. C represents a control reaction in which no acceptor/RT/NC is present.

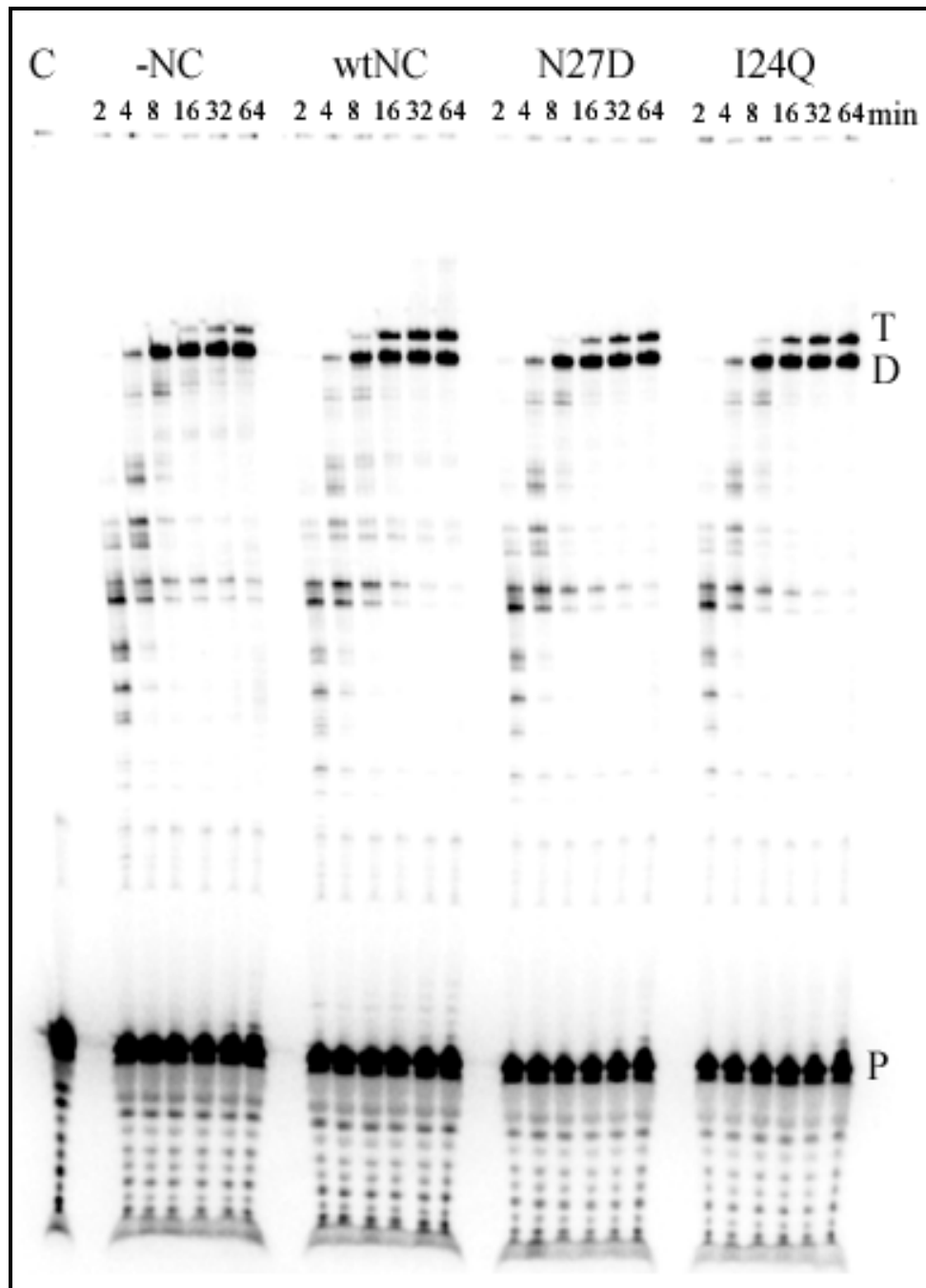


Figure 2-28: Autoradiogram of strand transfer assay with N27D and I24Q

Shown is an autoradiogram obtained from a strand transfer assay performed with N27D and I24Q NC mutants. The assay was carried out as described above.

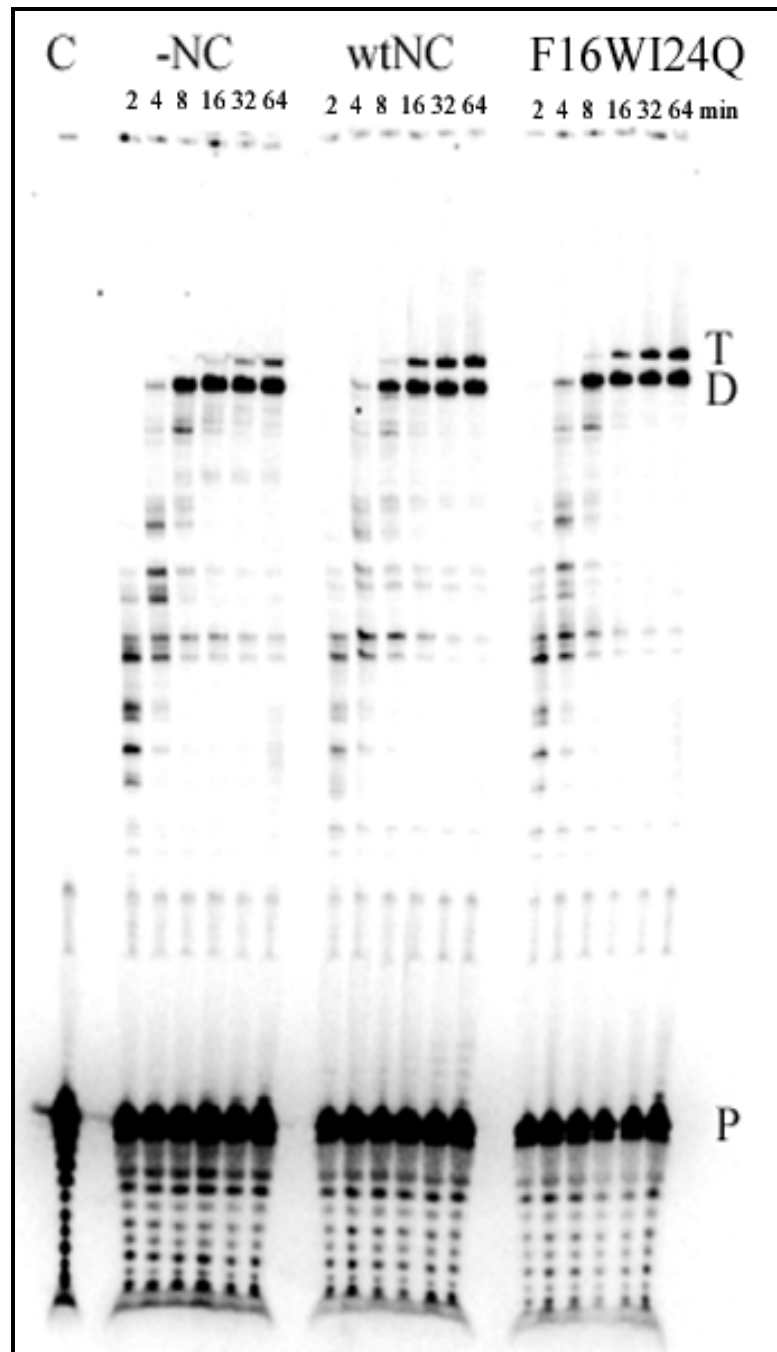


Figure 2-29: Autoradiogram of strand transfer assay with F16W/I24Q

Shown is the autoradiogram obtained from a strand transfer assay carried out with the F16W/I24Q NC mutant. The experiment was performed as described earlier.

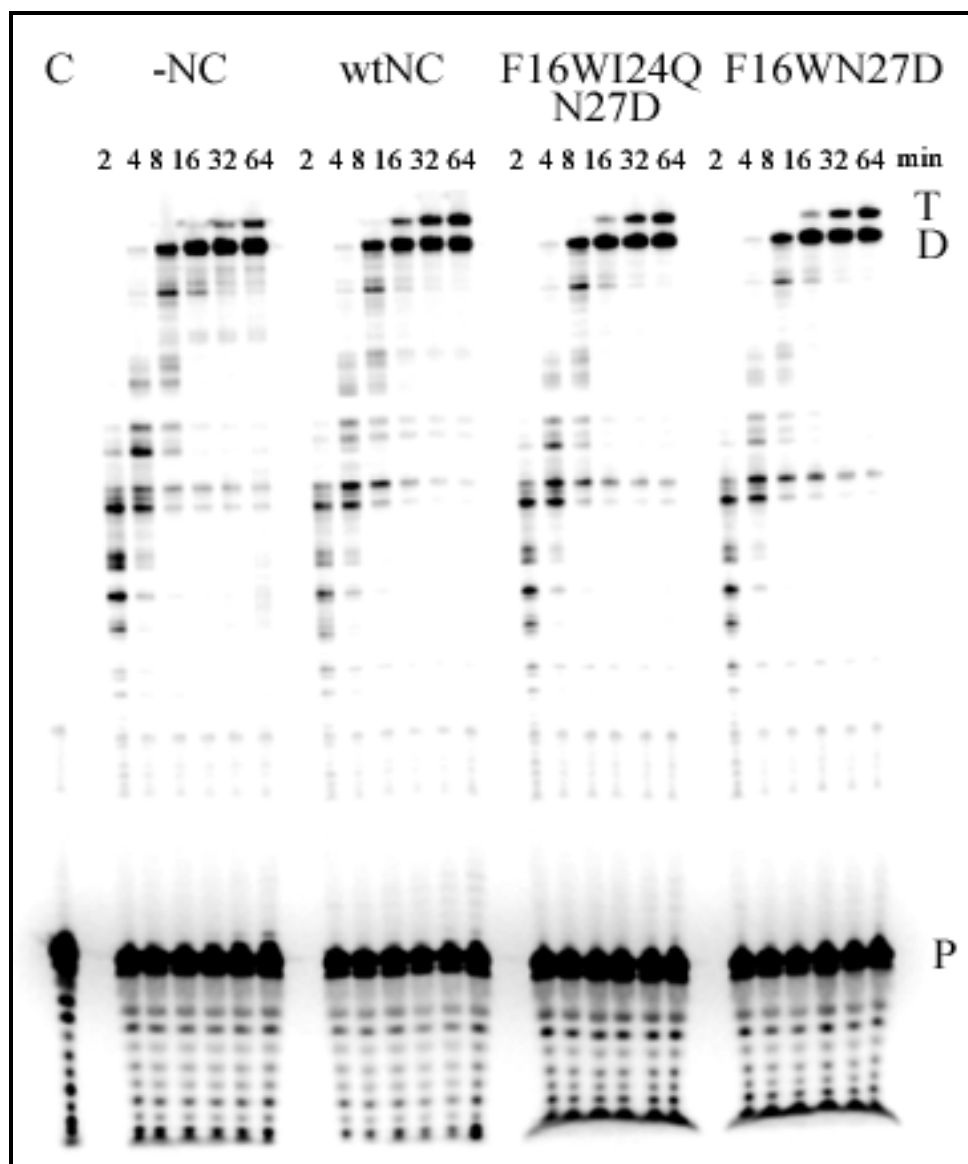


Figure 2-30: Autoradiogram of strand transfer assay with F16W/I24Q/N27D and F16W/N27D

Shown is the autoradiogram obtained from a strand transfer assay with F16W/I24Q/N27D and F16W/N27D NC mutants. The assay was carried out as described earlier.

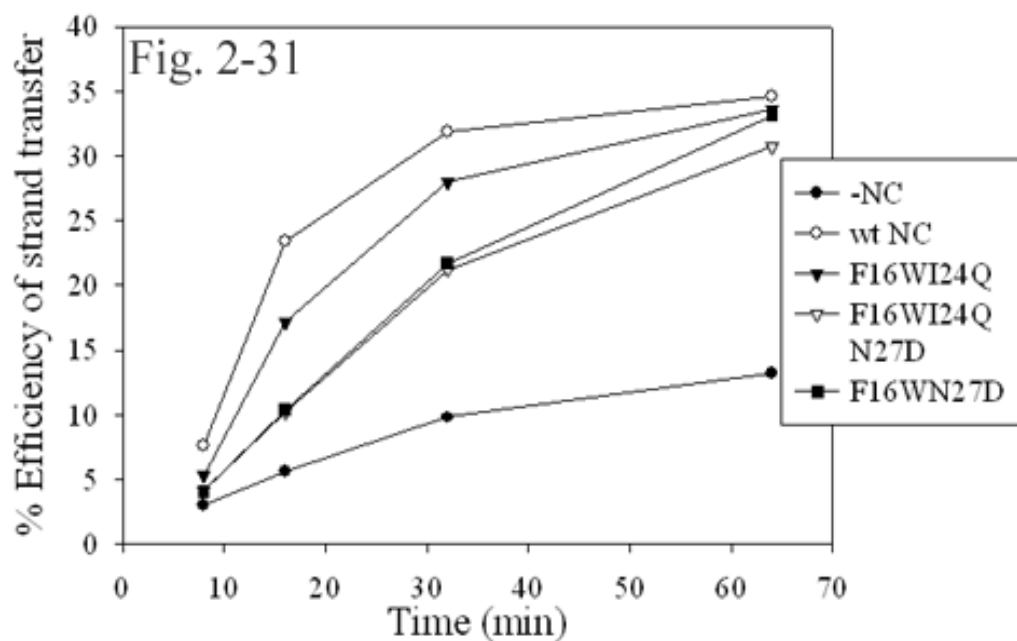


Figure 2-31: Graph showing efficiency of strand transfer *versus* time for F16W/I24Q, F16W/I24Q/N27D and F16W/N27D

Graph of efficiency of strand transfer *versus* time for F16W/I24Q, F16W/N27D and F16W/I24Q/N27D NC mutants is shown. The graph was made from quantification of the experiments shown earlier. The % efficiency of strand transfer was obtained using the following formula: $\text{Transfer DNA products (T)} / \{\text{Transfer DNA products (T)} + \text{Full length donor-directed DNA products (D)}\} * 100$, or $[\text{T} / (\text{T} + \text{D})] * 100$. It was then plotted against time in min.

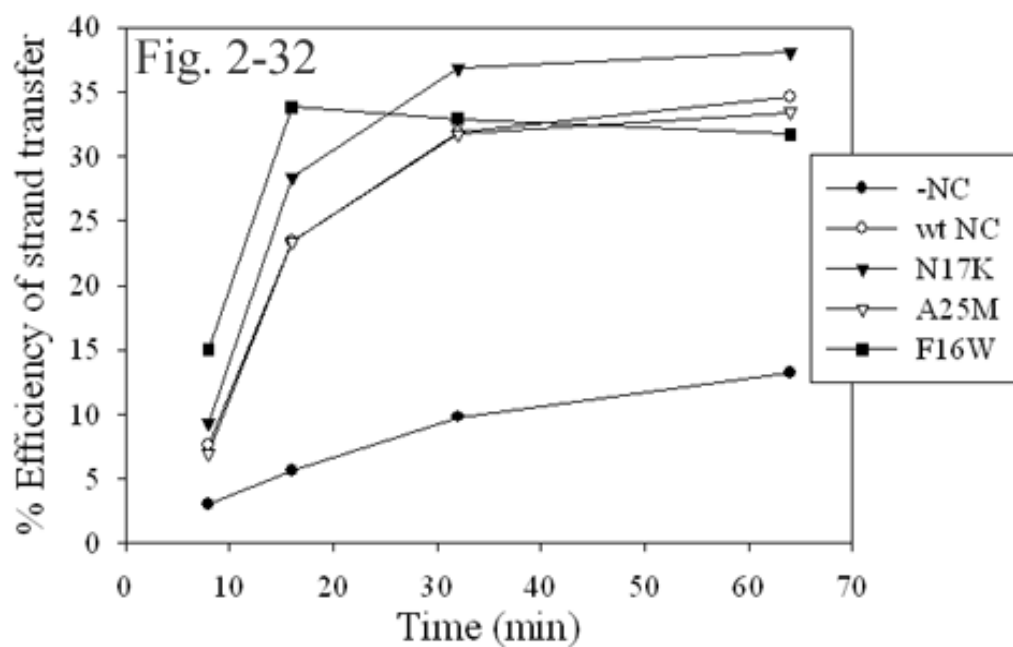


Figure 2-32: Graph showing efficiency of strand transfer *versus* time for N17K, A25M and F16W

Graph of efficiency of strand transfer *versus* time for N17K, A25M and F16W NC mutants is shown. The graph was plotted as described earlier.

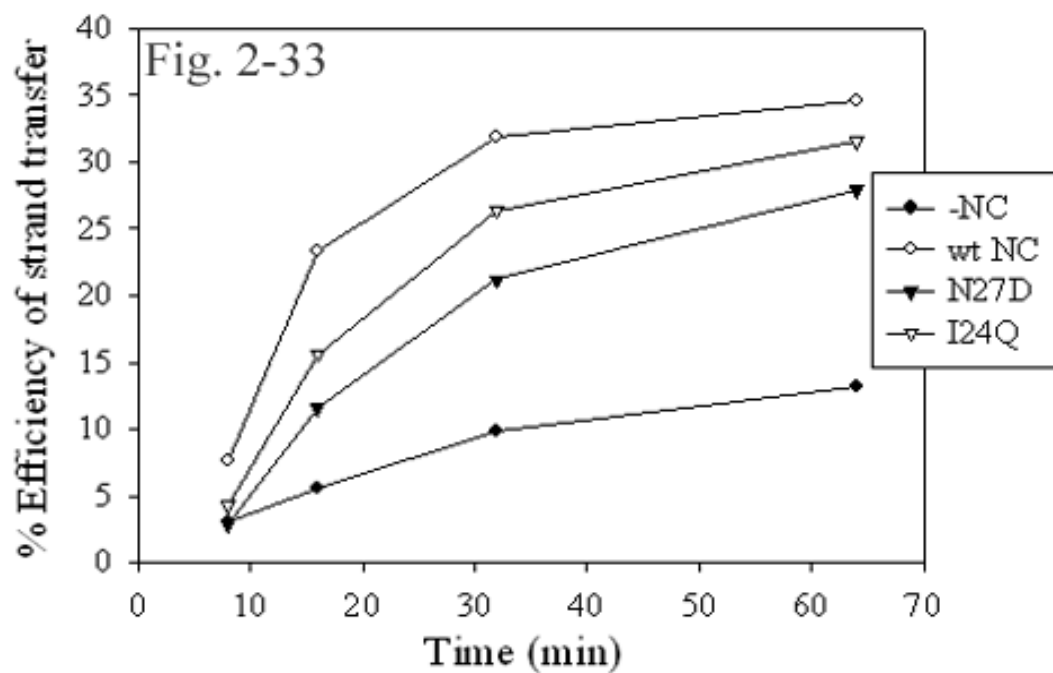


Figure 2-33: Graph showing efficiency of strand transfer *versus* time for N27D and I24Q

Graph of efficiency of strand transfer *versus* time for N27D and I24Q NC mutants is shown. The graph was plotted as described earlier.

Table 2-5
Sequence of forward and reverse primers used for PCR amplification of the *gag-pol* region

Region	Primer sequences
<i>GagPol</i> donor	a. 5' gatttaggtgacactatag <i>atata</i> gaaatgtggaaagga 3'
	b. 5' ttgtgtctctaccccagac 3'
<i>GagPol</i> acceptor	a. 5' gatttaggtgacactatag <i>tatat</i> cccctaggaaaaagggctgt 3'
	b. 5' ctgaagctctcttctggtgg 3'

2.7 Discussion

In this report we attempt to understand at an amino acid level why zinc finger one is more important than finger two to the helix destabilizing activity of HIV-1 NC (See Sec 2-1). The approach used was to make mutations in finger one using the corresponding amino acids in finger two (See Fig. 1-7). Five of the 14 total amino acids in each finger are different. Of the 5 changes, three of them represent a significant chemical alteration including two neutral to charge changes: (finger one to two) N17 to K38 and N27 to D48, and a hydrophobic to polar change: I24 to Q45. The other two are amino acids of similar hydrophobicity with those in finger 2 being more bulky: F16 to W37 and A24 to M46. Overall the five changes result in finger one being considerably more hydrophobic than two. Results showed that only I24Q and N27D mutations significantly decreased NC's helix destabilizing activity while A25M caused a small reduction. The N17K mutation had no measurable effect while F16W imparted greater helix destabilizing activity on NC. Double and triple mutant combinations with F16W combined with I24Q and N27D showed that the latter two were "dominant negatives" in that all the combination mutants were highly defective in both binding DNA and helix destabilization.

The simple approach of drawing conclusions from altering just these 5 amino acids partly ignores potentially important context effects. For example, NMR studies have shown that several interactions occur between finger one amino acids and other parts of the NC protein during binding to specific HIV stem-loop structures [77], [124], [44]. These include for example, F16 and N17 interacting with W37, I24 with valine and phenylalanine residues at positions 13 and 6, respectively, and N17 interacting with

proline and lysine at positions 31 and 33. Also, recent results implicate finger two in helix destabilizing as a W37 to leucine mutation in finger two greatly decreased NC binding to nucleic acids and helix destabilizing activity on the HIV TAR region [138]. Clearly, several interactions between finger one and two, and finger one and non-finger amino acids in NC are important to its biological functions, however, results suggest they may be less important to chaperone and helix destabilizing activity. Experiments show that finger mutants without an active finger two (1.1 or SSHS2) are still potent chaperone proteins *in vitro* and retain nearly wt NC helix-destabilizing activity (See Sec 2-1). In addition, truncated NC proteins missing the first 12 N-terminal non-finger amino acids (NC 12-55) showed strong helix-destabilizing activity. A double switch mutant in which F16 is replaced by tryptophan and W37 by phenylalanine retained most of its binding and helix destabilizing activity on TAR implying that the aromatic nature of the amino acids rather than a specific residue at a specific position is important to activity [138]. This is further supported by the high activity of NC 1.1 which has phenylalanine at positions 16 and 37.

The genetic flexibility of NC is further illustrated by cell culture experiments using mutated viruses. In one report, replacement of finger one with 7 different zinc fingers from a family of human cellular nucleic acid binding proteins (CNBPs) resulted in infectious viruses in 6 of the 7 cases [139]. In another report, alanine scanning mutagenesis was performed on finger one [140]; excluding F16 and the four amino acids that are part of the zinc binding motif (C15, C18, H23, and C28), alanine mutagenesis of the other 9 amino acids (in one case alanine 25 was changed to glycine) produced viruses with essentially wt NC infectivity that in some cases showed mild to moderate RNA

packaging defects. Although F16A was defective, an F16W mutant was comparable to wt virus, a result that correlated with our finding that F16W is as good as or better than wt in helix destabilization and strand transfer. In the F16W mutant, phenylalanine gets replaced with a more bulky amino acid namely tryptophan. Perhaps this enables it to intercalate between nucleic acid bases even better, thereby resulting in enhanced chaperone activity. This could also explain why A25M exhibits similar or slightly diminished unwinding activity with structured substrates in comparison to wt NC, since methionine is also a more bulky group than alanine. Also interesting with respect to our results was the finding that N27 could be changed to alanine with little effect on infectivity. Although this position is fairly highly conserved, HIV isolated with serine, histidine, isoleucine, and tyrosine at this position have been reported (HIV genome sequence information was obtained from the HIV sequence database provided by Los Alamos National Laboratory (<http://hiv-web.lanl.gov/content/hiv-db/mainpage.html>)). The N27D mutant in our studies was quite defective in DNA binding and helix destabilization, but this is a much more drastic change than those changes noted above. The NC protein of the virus used in the above study had a threonine at position 24 rather than isoleucine as is present in the strain we used, but the fact that it could be changed to alanine without loss of infectivity suggests some flexibility at this position. This position is not as strongly conserved as the other 4 examined in our assays. Many HIV isolates have leucine at this position and some have threonine and valine. The I24Q and an I24E (glutamic acid) mutant that we also tested (Data not shown) were quite defective in helix destabilization and the latter bound DNA poorly. Perhaps there is a limit to how polar the amino acid at this position can be. Isoleucine 24 along with V13, F16, and A25 have been implicated as part of an

important hydrophobic cleft in finger one that is required for binding and destabilizing nucleic acids [107]. The fact that some virus strains have threonine at position 24 and that alanine can be substituted without loss of viability suggests that the amino acid at this position need not be highly hydrophobic. It is also possible that NC proteins with reduced helix destabilizing activity may be functional during infection. The NC protein of Moloney murine leukemia virus (MuLV) has very little helix destabilizing activity in *in vitro* assays [130] and it seems unlikely that the 6 functional fingers from the human CNBP proteins (See above) would all have high destabilizing activity. Therefore it is possible that HIV-1 NC could get by with less, as long as other important NC functions were not compromised. We are in the process of testing the mutants made in these experiments in the context of cellular infections. Preliminary experiments suggest that there is no defect in replication with N27D mutant viruses in H9 cells (Data not shown); supporting the hypothesis that reduced helix destabilizing activity is not detrimental to virus fitness. It would also be interesting to test the helix destabilizing activity of viable mutations at position 24 as well.

The N17K mutation tested here has also been examined in cell culture infections [141]. Interestingly, viruses with this mutation show no replication defects and actually have a 7-9 fold increased transduction frequency compared to wt NC. In addition to packaging viral RNA better than wt NC, there was also an increased packaging of non-viral RNA. Several other mutations in the finger amino acids that increased net positive charge of the fingers also lead to viable viruses with transduction levels as good as or better than wt NC. The N17K mutation was essentially equivalent to the wt NC in all the assays we tested it in so it is not surprising that it would function well. Despite this,

BLAST searches recovered no isolated viruses with N17K mutations, suggesting that this change is selected against *in vivo* [141].

The strand transfer assays were reasonably consistent with helix-destabilization assays in that those mutants defective in the latter generally showed some decrease in strand transfer. Mutant N27D and the double and triple mutants (F16W/N27D and F16W/I24Q/N27D) containing this mutation appeared to be the most defective while I24Q and F16W/I24Q were moderately less stimulatory than wt NC (See Sec 2-6). The other point mutants were essentially the same as wt NC with N17K showing a small enhancement. The *gag-pol* genome region used in the assays is highly structured [135] and it seems surprising that all the mutants, even the double and triple mutants and N27D which were highly defective in helix-destabilization and DNA binding, showed a clear enhancement. However, previous assays with finger mutant 2.2, which is highly defective in helix destabilization assays (Tables 2-2, 2-4 and [91]), showed stimulation of strand transfer on the *gag-pol* substrate, albeit at a reduced level compared to wt NC. In addition, 2.2 was essentially identical to wt for stimulation on a low-structured substrate [135]. These results suggest that the aggregation/condensation activity of NC, which all the mutants appeared to retain (See Sec 2-4), is a major driving force for strand transfer while helix destabilization may be less important, especially in low structured genome regions.

From our results, the most important amino acid differences between finger one and two with respect to chaperone activity are I24Q and N27D. Both are extreme chemical changes that strongly decrease the hydrophobic nature of finger one. The overall more hydrophobic nature of finger one could be important to its apparent

advantage for helix destabilization. On a mechanistic level, mutant N27D caused a large defect in nucleic acid binding and this alone could have been responsible for its low activity. Low binding may have occurred from repulsion of nucleic acids by the additional negative charge of the aspartic acid side chain or a dramatic change in protein folding. The latter is unlikely since viral clones with this mutation showed no defect in replication in H9 cells (See above). Mutant I24Q did not significantly alter nucleic acid binding. This could be due to the neutral nature of glutamine which probably did not alter the charge density at that position. This suggests that no deleterious change in overall structure had occurred. However, perhaps by decreasing the hydrophobicity of finger one, substitution of glutamine for isoleucine reduces NC's ability to interfere with hydrophobic base stacking. This would be consistent with the previously proposed role of the hydrophobic pocket of finger one (See above, [138])

Chapter 3 Analysis of the nucleic acid chaperone activity of other retroviral NC proteins

3.1 Introduction

Retroviral nucleocapsid (NC) proteins are very small, basic proteins that coat the genomic RNA. They act as molecular chaperones during several steps in the viral replication cycle. They catalyze transitional unfolding and re-pairing of nucleic acids such that the optimal conformation with maximum number of base pairs is attained. Previous chapters dealt with in-depth analysis of nucleic acid chaperone activity of HIV-1 NC. Work done in this chapter sheds light onto helix-destabilizing activity of other retroviral NC proteins. These include those from Moloney murine leukemia virus (MuLV NC), the Mne strain of Simian immunodeficiency virus (SIV NC_{p8}), HIV-1 NC_{p15} and HIV-1 NC_{p9}. HIV-1 NC_{p15} is first produced as a result of proteolytic cleavage from the *Gag* precursor polyprotein. NC_{p9} results due to further cleavage from NC_{p15}. This is further cleaved to just under 7 kDa to yield the functional NC protein used in previous chapters [1].

Table 3-1 lists amino acid sequences of the various NC proteins used in this study. SIV belongs to the *Lentivirus* genus of retroviruses. The nucleocapsid protein of SIV (SIV NC_{p8}) is 52 amino acids in length and has two zinc fingers. Each zinc finger is 14 amino acids in length and coordinates one Zn²⁺ ion. Each finger has a Trp residue. A wealth of information about several properties of SIV NC_{p8} has been obtained through mutational studies and *in vitro* assays. Finger mutations in SIV NC_{p8} have revealed that lack of zinc coordination affects its stability and interferes with processing of the *Gag*

precursor polyprotein [142]. In another vaccine study, a group of *Macaca nemestrina* primates were first infected with a pathogenic wt strain of SIV NCp8. This was then followed by immunization with DNA that produces replication defective yet structurally complete virions. These virions were obtained from NC deletion mutant clones and contained a full set of processed viral proteins capable of completing nearly all the steps during one viral replication cycle. This appears to offer protection or at least delay progression onto AIDS [143]. In one other study, four amino acids were deleted in the carboxyl terminal zinc finger of SIV NCp8, producing non-infectious viruses (though capable of completing most of the steps in viral replication). Immunization with DNA expressing this NC mutant provided substantial humoral immunity [144]. In one experimental study with macaques, no antibody production/disease progression was observed. Gorelick *et al.* have demonstrated that mutations in both zinc fingers of SIV Mne NC produce replication defective virions both *in vitro* and *in vivo*. The basic residues found in SIV NC have been shown to be very crucial for incorporating genomic RNA. However, RNA encapsidation in SIV was found to depend more on finger two rather than finger one as is the case in HIV-1 [145]. Optically detected magnetic resonance (ODMR) studies have revealed that both the zinc fingers of SIV NC_{p8} bind simultaneously to oligonucleotides, and were found to contribute to its binding free energy, though not additively. Significant quenching of fluorescence from both Trp residues suggests that hydrophobic stacking occurs between the indole ring and nucleobases. The binding order of preference was similar to that of HIV-1 NC namely: G ~ I (Inosine) > T > U > C > A. SIV NC_{p8} bound to alternating bases in a sequence-specific manner similar to HIV-1 NC [146]. NMR studies conducted with the zinc finger

domains (amino acids 13-51) of SIV NC_{p8} have revealed that the main structural differences between SIV NC_{p8} and HIV-1 NC occur at the flexible linker and not the zinc fingers. However, orientation in which the two proteins interact with oligonucleotides was found to be quite similar [147].

Murine leukemia virus belongs to the gammaretrovirus family. MuLV NC is 60 amino acids in length and is located at the C-terminus of *Gag*. It is produced as a result of proteolytic cleavage from *Gag* upon maturation. It contains only one zinc finger motif. Mutational studies with MuLV NC have revealed that it stimulates DNA synthesis by RT at pause sites [148]. Experiments done with MuLV NC zinc finger mutants CCCC or CCHH have shown that these mutations do not affect RNA recognition, t-RNA primer placement, RNA packaging and dimer RNA structure. However, they do seem to greatly affect cDNA synthesis and hence produces non-infectious virions defective in replication [149], [150]. Further, NC native zinc finger architecture was found to be critical for replication events preceding integration and possibly viral integration itself. In one study, when all the cysteine residues were replaced with serines, RNA packaging was greatly reduced. Gorelick *et al.* also point out that this sequence is fairly conserved across retroviral NC proteins and could be related evolutionarily to the zinc fingers that are involved in specific sequence recognition in double stranded DNA [79]. Further, Rein *et al.* have also shown through mutational studies that a central region of MuLV NC consisting of residues 16 through 23 is critical for RNA packaging. This region is followed by the zinc finger's cysteine array and flanked by basic residues. Further, the basic region and zinc finger have also been demonstrated to be critical for assembly and release of virus particles [151]. D'Souza *et al.* have shown that unlike HIV-1 NC, the

binding affinity of MuLV NC to the stem loops found in its genomic RNA is quite low [152]. Allain *et al.* have shown that the basic residues and not the zinc finger in MuLV NC is critical for enhancing the minus strand transfer step during reverse transcription [153]. Single molecule stretching experiments and mutational studies have been carried out to evaluate the chaperone activity of HIV-1 and MuLV NC. Results have shown that HIV-1 NC but not MuLV NC lowers the cooperativity of DNA helix-coil transitions. Since this is a measure of helix-destabilizing activity it suggests that HIV-1 NC has more potent destabilizing activity. Also consistent with this finding, Williams *et al.* point out that a specific two finger architecture is most likely required for unwinding and hence optimal chaperone activity during viral reverse transcription. Overall the findings suggest that MuLV can efficiently replicate with an NC protein possessing little helix-destabilizing activity in comparison to HIV-1 [130].

Finally, compounds that oxidize sulfur atoms by inducing disulfide cross linking in cell-free MuLV virions have been identified [154]. This effect has also been observed in HIV-1 but not in HFV virions (Human Foamy Virus) [155]. It should be noted that HFV belongs to the *Spumaretrovirus* genus of retroviruses which lack zinc fingers. Thus, the zinc finger of MuLV NC forms the primary target for these anti-retroviral compounds. Other zinc ejecting compounds that potentially inhibit HIV-1 through NC deactivation are also being tested [155]. Further studies on NC proteins could potentially lead to more and better inhibitors for use as potential therapeutics.

3.2 Materials

DNA oligonucleotides for the fluorescence resonance energy-transfer (FRET) assays were purchased from Integrated DNA Technologies (Coralville, IA). See Chapter 2 for a list of other materials used. All of the chemicals were purchased from Sigma Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA).

3.3 Methods

FRET assay to detect DNA/DNA annealing: Refer to Chapter 2 for details.

Preparation of HIV-1 wt NC: Wild-type HIV-1 NC from either the MN strain (GenBank accession number: M17449), the ARV strain or pNL4-3 was used in this study. Wild-type MN NC was expressed and purified as described [136]. The construct that expresses wild-type ARV NC (GenBank accession number: K02007) was graciously provided by Dr. Charles McHenry (University of Colorado) and this protein was prepared as described previously [107]. Wild type NC protein from the pNL4-3 sequence (GenBank accession number AF324493) was prepared essentially as described [75]. The three wt NC proteins differ by no more than five amino acids, which are all functionally conserved (See Table 3-1). SIV NCp8, MuLV NC, NCp9 and NCp15 proteins were kindly provided by Dr. Robert J. Gorelick, SAIC, Frederick, MD. NC aliquots were stored at -80°C in 50 mM Tris-HCl (pH 7.5), 10% glycerol, and 5 mM 2-mercaptoethanol.

3.4 Results

Annealing assays performed with wt NC and other retroviral NC proteins: FRET annealing assays were performed as described in Chapter 2 with the 0.0, 5.8 and 9.0dna substrates. However, while the assay was carried out for 4 min as before with the 5.8dna, the annealing was monitored for only 1 min with the 0.0dna. This was done due to rapid annealing kinetics of the non-structured substrate as most of the annealing was completed within the first min of the time course reaction. Five retroviral NC proteins were tested for their annealing activity. These include MuLV NC, SIV NC_{p8}, HIV NC_{p9}, HIV NC_{p15} and wt HIV-1 NC (NC_{p7}). Figure 3-1 and Table 3-2 summarize the results obtained from annealing assays carried out with 0.0dna. Rate constants for complement annealing were calculated as described earlier. With the unstructured substrate, complementary nucleotides annealed very rapidly even in the absence of NC, and NC clearly enhanced annealing even further resulting in about a 2-fold increase in the rate constant. This demonstrates NC's ability to enhance annealing even in the absence of secondary structure, presumably by aggregation/condensation. While MuLV NC and NC_{p9} showed no stimulation on 0.0dna, SIV NC_{p8} appeared to retain some aggregation/condensation activity based on this assay. NC_{p15} stimulated annealing close to wt NC levels. This is consistent with the idea that this activity results mostly from the highly positively charged NC backbone amino acids that act to neutralize negative charges on the phosphate backbone of nucleic acids.

Figure 3-2 and Table 3-3 show the effect observed with various NC proteins on 5.8dna. On this more strongly folded substrate clear groupings emerged. NC_{p15} was able

to stimulate annealing as well as wt NC. SIV NC_p8 showed reduced annealing in comparison to NC_p15, but was clearly better than MuLV NC. MuLV NC showed about a 60-70% rate reduction in comparison to wt HIV-1 NC. Interestingly, NC_p9 facilitated annealing was nearly 2-fold greater than wt HIV-1 NC (Fig. 3-2 and Table 3-3). The same general pattern was observed with the strongest structure, 9.0dna (Fig. 3-3 and Table 3-4). In this case annealing was much slower than with 5.8dna (approximately 4-7 fold) reflecting greater stability of the substrate; and the assay was performed over 16 min rather than 4 min as with 5.8dna. Very little annealing was observed in reactions without NC and no rate values were obtained. Once again NC_p15 was similar to wt HIV-1 NC. SIV NC_p8 showed some reduction in annealing. No detectable stimulation was observed with MuLV NC. Again, NC_p9 stimulated annealing better than wt HIV-1 NC showing about a 2-fold increase (Fig. 3-3). Shown in Tables 3-3 and 3-4 are the rate constants for NC proteins on structured DNA substrates (5.8dna and 9.0dna). Annealing assays with SIV NC_p8 and MuLV NC on 9.0dna were also carried out using reduced salt conditions. As expected, the rate of annealing increased by nearly 2-fold at 40 mM KCl and even higher; by nearly 4-fold at 20 mM with SIV NC_p8. The effect of low salt buffer was even more pronounced in case of MuLV NC. The observed rate constant increased by nearly 4-fold and 7-fold at 40 mM and 20 mM KCl, respectively (Table 3-4).

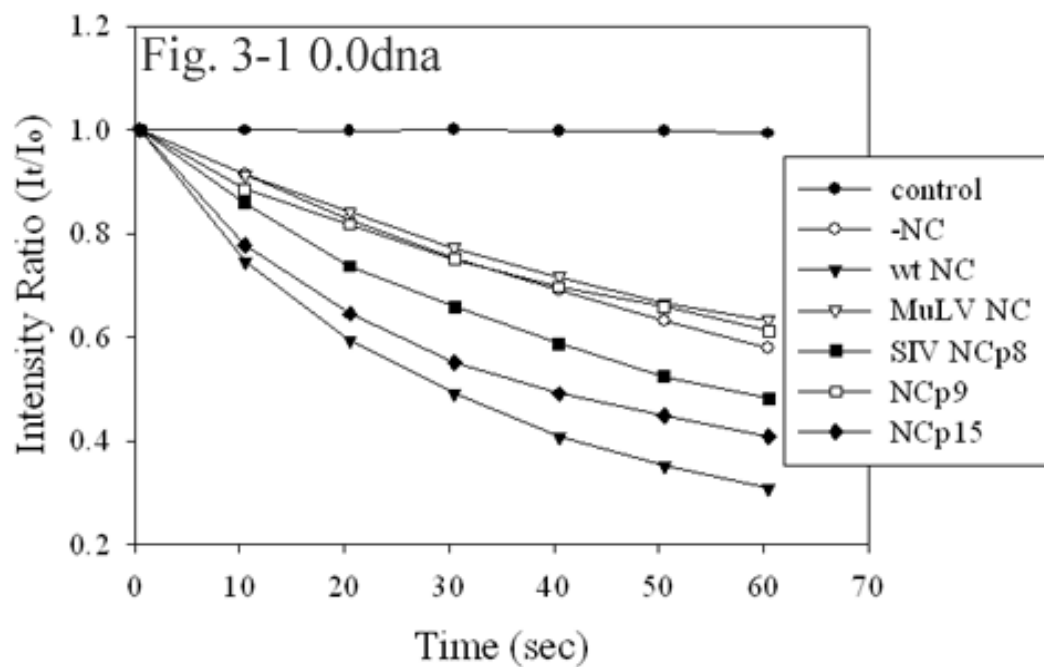


Figure 3-1: FRET assay with retroviral NC proteins on 0.0dna

Shown above is the FRET assay performed with the five retroviral NC proteins on 0.0dna. C denotes a control reaction in which no complementary DABCYL DNA was present. -NC denotes a reaction in which no NC was present. The assay was carried out as described previously except that the annealing reaction was monitored for only 1 min (See Chapter 2).

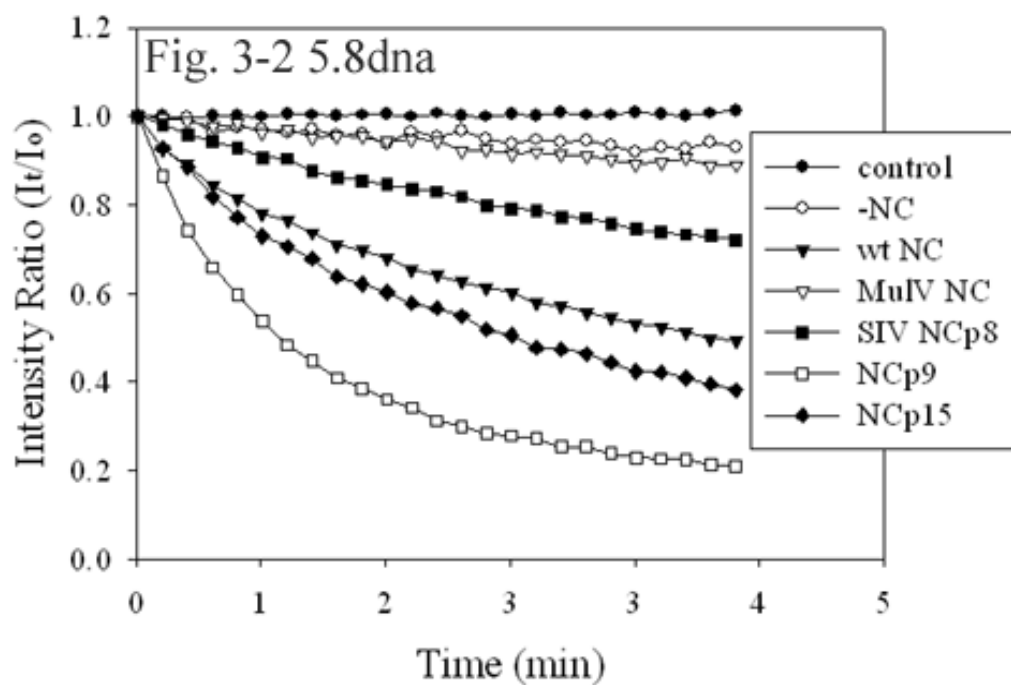


Figure 3-2: FRET assay with retroviral NC proteins on 5.8dna

Shown above is the FRET assay performed with the five retroviral NC proteins on 5.8dna. C denotes a control reaction in which no complementary DABCYL DNA was present. -NC denotes a reaction in which no NC was present. The assay was carried out as described previously (See Chapter 2).

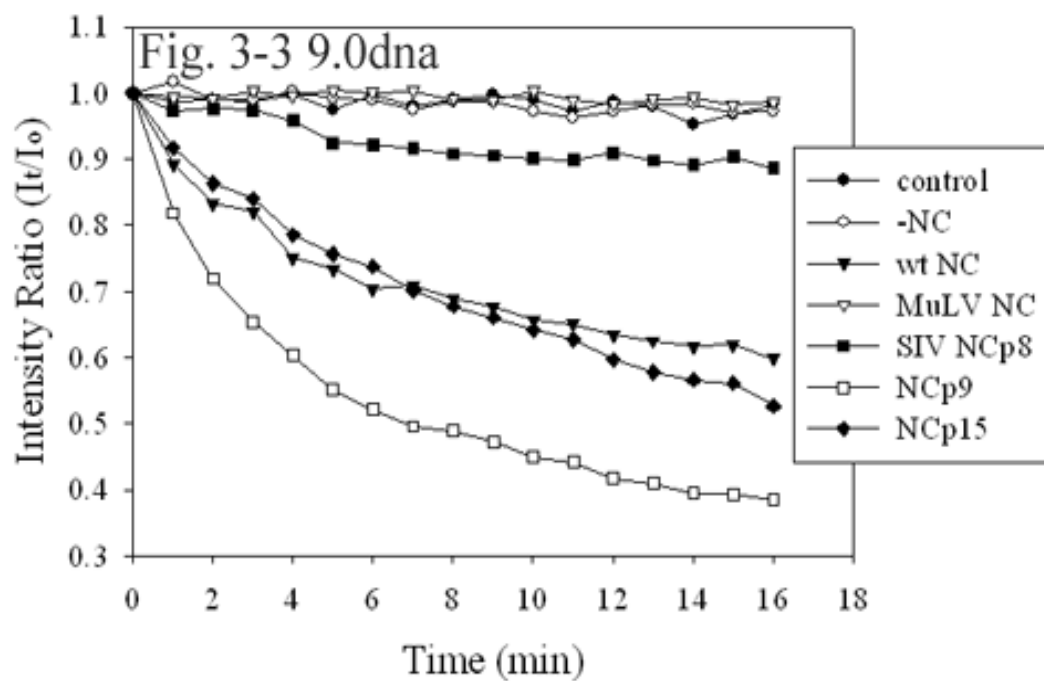


Figure 3-3: FRET assay with retroviral NC proteins on 9.0dna

Shown above is the FRET assay performed with the five retroviral NC proteins on 9.0dna. C denotes a control reaction in which no complementary DABCYL DNA was present. -NC denotes a reaction in which no NC was present. The assay was carried out as described previously (See Chapter 2).

Table 3-1
Amino acid sequences of various NC proteins used in this study

¹VIRUS SOURCE	AMINO ACID SEQUENCE
HIV - 1 MN	MQRGNFRNQRK II KCFNCGKEGHIAKNCRAPRK RG CWKCGKEGHQMKDC TERQAN
HIV - 1 ARV	MQRGNFRNQRKTVK CF NCGKEGHIAKNCRAPRKK GC WRCGREGHQMKDC TERQAN
HIV - 1 NL4 - 3 (NCp7)	IQ KGNFRNQRKTVK CF NCGKEGHIAKNCRAPRKK GC CWKCGKEGHQMKDC TERQAN
HIV - 1 NL4 - 3 (NCp9)	IQKGNFRNQRKTVK CF NCGKEGHIAKNCRAPRKK GC CWKCGKEGHQMKDC TERQANFLGKIWPSHKGRPGNFL
HIV - 1 NL4 - 3 (NCp15)	IQKGNFRNQRKTVK CF NCGKEGHIAKNCRAPRKK GC CWKCGKEGHQMKDC TERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPEESFRFGEETTTPSQK QEPIDKELYPLASLRSFLGSDPSSQ
MuLV NC	ATVVSGQKQDRQGGERRRSQ ¹ LRDQ CAYCKEKGHWAKDC KKPRGPRGPR PQTSL
SIV NCp8	AQQKGPRKPIK CW NCGKEGHSARQCRTPRRQ GC WKCGQMGHVMAK CP DR QAG
<p>1- The three wt HIV-1 NCp7 proteins used in this assay differ by no more than five amino acids (highlighted in yellow), which are all functionally conserved. Amino acid sequences of other retroviral proteins namely SIV NCp8, MuLV NC, NCp9 and NCp15 proteins are also shown. Finger domains of all NC sequences are highlighted in pink.</p>	

Table 3-2. Rate constant calculation for NC proteins on 0.0dna substrate		
¹ Name	² t_{1/2}(min)	² k*10⁻³(min⁻¹)
-NC	1.27	0.549 ± 0.046
wt NC	0.598	1.16 ± 0.049
MuLV NC	1.53	0.469 ± 0.119
SIV NCp8	0.956	0.732 ± 0.102
NCp9	1.47	0.478 ± 0.077
NCp15	0.797	0.870 ± 0.024
<p>1-HIV strain MN NC was used in these assays and was prepared essentially as described (See Methods). All of the other retroviral proteins were a kind gift of Dr. Robert J. Gorelick, SAIC, Frederick, MD.</p> <p>2-<i>k</i> (rate constant) values were calculated from the t_{1/2} values by dividing 0.693 by t_{1/2} as described. Results are an average of 2-3 experiments ± standard deviations (shown for <i>k</i> only).</p>		

Table 3-3. Rate constant calculation for NC proteins on 5.8dna substrate		
¹ Name	² t_{1/2}(min)	² k*10⁻³(min⁻¹)
wt NC	4.31	0.161 ± 0.005
MuLV NC	23.3	0.030 ± 0.005
SIV NCp8	8.67	0.080 ± 0.009
NCp9	1.90	0.364 ± 0.006
NCp15	3.06	0.229 ± 0.029
<p>1-HIV strain MN NC was used in these assays and was prepared essentially as described (See Methods). All of the other retroviral proteins were a kind gift of Dr. Robert J. Gorelick, SAIC, Frederick, MD.</p> <p>2-<i>k</i> (rate constant) values were calculated from the t_{1/2} values by dividing 0.693 by t_{1/2} as described. Results are an average of 2-3 experiments ± standard deviations (shown for <i>k</i> only).</p>		

Table 3-4. Rate constant calculation for NC proteins on 9.0dna substrate		
¹ Name	² t_{1/2}(min)	² k (min⁻¹)
wt NC	29.2	0.024 ± 0.001
MuLV NC	834	0.001 ± 0.001
SIV NCp8	42.7	0.016 ± 0.002
NCp9	13.6	0.052 ± 0.010
NCp15	19.4	0.037 ± 0.009
MuLV NC (20 mM KCl)	151	0.005
MuLV NC (40 mM KCl)	197	0.004
SIV NCp8 (20 mM KCl)	9.80	0.071
SIV NCp8 (40 mM KCl)	17.9	0.039
<p>1-HIV strain MN NC was used in these assays and was prepared essentially as described (See Methods). All of the other retroviral proteins were a kind gift of Dr. Robert J. Gorelick, SAIC, Frederick, MD. 2-k (rate constant) values were calculated from the t_{1/2} values by dividing 0.693 by t_{1/2} as described. Results are an average of 2-3 experiments ± standard deviations (shown for <i>k</i> only for 80 mM KCl assays).</p>		

3.5 Discussion

In this work, we wished to analyze the helix-destabilizing activity of other retroviral NC proteins namely SIV NC_{p8} and MuLV NC, and two precursor forms of wt HIV-1 NC; NC_{p9} and NC_{p15}. MuLV NC has only one zinc finger motif. Moreover, it does not bind tightly to RNA stem loops under *in vitro* conditions [152]. Single molecule stretching experiments have shown that it does not lower the cooperativity of DNA helix-coil transitions, even at saturating levels, whereas HIV-1 NC exhibits optimal chaperone activity at a much lower concentration. Also Williams *et al.* point out that the differences between the two proteins could be due to differential binding affinities with HIV-1 NC binding much more tightly. Further, it is also possible that MuLV NC does not require unwinding activity for stimulating minus-strand transfer under *in vivo* conditions, while HIV-1 NC does. *In vitro* strand transfer assays conducted using MuLV NC in the HIV-1 R region and DNA stretching studies conducted by Williams *et al.* augment this hypothesis. However, it should be noted that the secondary structure element that has to be unwound in order to facilitate MuLV minus strand transfer is much less stable than the highly structured HIV-1 TAR region. Therefore MuLV may not need as much destabilizing activity to carry out this process [130].

MuLV NC was found to be highly defective in helix-destabilization in comparison to wt NC in our assays (Figs. 3-2, 3-3 and Tables 3-3, 3-4). This correlates with the findings of Williams *et al.* described above [130]. It is worthwhile to note that the single finger of MuLV NC more closely resembles finger two of HIV-1 NC in that it has an aspartic acid residue in the same location as finger two of HIV-1 NC (See Fig. 3-1). Earlier helix-destabilization studies conducted with HIV-1 NC showed that replacing

asparagine with this residue was severely detrimental to its helix-destabilizing activity. It should also be noted that the N27D mutant was defective in nucleic acid binding and in *in vitro* strand transfer assays (See Chapter 2). It is reasonable to hypothesize that this amino acid residue could be partly responsible for MuLV NC's diminished nucleic acid binding and hence unwinding properties. Annealing assays conducted with non-structured and structured substrates show that MuLV NC is highly defective in both aggregation and unwinding.

SIV NC_{p8} exhibits reduced annealing as well; however it is clearly better than MuLV NC. The difference in annealing rates between SIV NC_{p8} and wt HIV-1 NC was nearly two fold (See Tables 3-3 and 3-4). It can be seen that SIV NC_{p8} shares high homology with HIV-1 NC especially in the zinc finger residues (See Table 3-1). Once again, it did not exhibit any defects in nucleic acid aggregation (See Table 3-2).

Further, while NC_{p15} enhances annealing close to wt HIV-1 NC levels, NC_{p9} was found to be better than wt HIV-1 NC (See Tables 3-3 and 3-4). The observed differences between NC_{p15} and NC_{p9} could be due to the presence of additional glutamine and glutamic acid residues in NC_{p15} (See Table 3-1). The isoelectric point (pI) values for wt HIV-1 NC, NC_{p9} and NC_{p15} were found out to be 10.08, 10.19 and 9.59 respectively (All pI values were obtained from ExPASy Proteomics Server of the Swiss Institute of Bioinformatics (SIB)). Upon comparing the pI values, it can be seen that NC_{p9} adds higher net positive charge. This could have minimized electrostatic repulsion between the phosphate groups of nucleic acids during annealing. Also, it should be noted

that NC_p9 was defective in the aggregation assay with 0.0dna. It is not clear what causes this apparent reduced aggregation activity of NC_p9. Clearly, further studies are needed to fully understand the various structure-function relationships of these proteins.

Finally, low salt (20 mM and 40 mM KCl) conditions were used to gain insight into the *duplex-stabilizing* properties of SIV NC_p8 and MuLV NC (See Tables 3-2 and 3-4). Optimal Mg²⁺ concentrations (6 mM) for reverse transcriptase assays were used in the above described experiments. As expected, under high salt conditions, higher screening of duplexes is brought about by high salt/multivalent cations. Thus, more effective annealing in the absence of a catalyst under high salt conditions most likely accounts for the weaker effect of NC enhanced annealing under low ionic strength conditions. The observed rate enhancement facilitated by NC is hence more pronounced under low ionic strength conditions, since there are less competing cations to displace NC [82]. A similar effect was observed in the annealing rates which increased by a huge factor, about 3-4 fold with lower Mg²⁺ (1 mM) concentration in the annealing assays performed earlier with HIV-1 NC mutants (See Chapter 2). However, the concentration of non-complexed Mg²⁺ in cells may be much lower [156] and results have shown that NC is more active with lower ionic strength [55]. This may also occur with MuLV NC and should be tested in the future.

Chapter 4 General Discussion

AIDS continues to be the one of the biggest threats facing mankind today. It has claimed more than 25 million human lives since its discovery in 1981 [6]. The ultimate goal of researchers worldwide is to develop efficient vaccines and drugs such that this virus can be eradicated. A thorough understanding of the virus and various steps in its replication cycle could help achieve this goal. HIV is one of the most recombinogenic of all retroviruses and has a high mutation rate. This has contributed to the rapid emergence of several subtypes and now intersubtype recombinant forms of the virus [5]. HIV-1 NC is a small, basic protein that coats the viral genome. The multitude of functions performed by this protein has caused it to gain equal importance on par with the three other classic retroviral proteins namely, reverse transcriptase (RT), integrase (IN) and protease. Thus, NC has emerged as a prime target for vaccine research and drug therapy. Obviously, it is of utmost importance to gain an extensive, in depth knowledge about NC; its structure, various properties and structure-function relationships. The approach used in this thesis of studying mutated proteins will also help define the genetic flexibility of NC. If NC is highly tolerant of mutations, as seems to be the case with RT, drug resistant mutants are likely to emerge during NC targeted therapy.

Work done in this thesis sheds light onto the chaperone activity of HIV-1 NC. Chaperones are proteins that can catalyze the rearrangement of nucleic acid molecules such that the most thermodynamically stable conformation is achieved [41]. Chaperone proteins are found in other organisms as well. These include gene 32 protein found in T4 bacteriophages (T4 gp32), single-strand binding protein in *E.coli* (*E.coli* SSB) and ribonucleoprotein A1 found in humans (hnRNP A1) (For a more complete list see

<http://mendel.imp.ac.at/home/Birgit.Eisenhaber/RNA-chaperones/list.html>). However, HIV-1 NC continues to be the most widely studied chaperone protein.

My thesis attempts to understand at an amino acid level why zinc finger one of HIV-1 NC is more important than finger two to its helix destabilizing activity (See Sec 2-1). Five of the 14 total amino acids in each finger are different. The approach used was to incrementally replace each differing residue in finger one by the one at the corresponding location in finger two (See Fig. 1-7). The most important amino acid differences between finger one and two with respect to chaperone activity are I24Q and N27D (See Chapter 2). Overall, NC is quite tolerant of single amino acid changes as even the most defective single point mutants (I24Q and N27D) retained about 50% helix destabilizing activity (See Sec 2-4) and nearly full aggregation activity (As judged by assays with the non-structured substrate). Double and triple mutants were more defective than wt NC but still retained measurable activity (See Sec 2-4). The high genetic flexibility of NC which is further illustrated by previous noted experiments showing that replacing the entire finger of NC with human protein zinc fingers in most cases did not abolish infectivity [139]. The unfortunate likelihood is that viable mutants would arise easily upon drug NC directed drug treatments. Still such drugs could be very useful in combination therapy with drugs targeting other proteins.

Also as discussed previously in Chapter 2, interesting with respect to our results was the finding that N27 could be changed to alanine with little effect on infectivity [140]. Although this position is highly conserved, HIV isolated with serine, histidine, isoleucine, and tyrosine at this position have been reported (HIV genome sequence information was obtained from the HIV sequence database provided by Los Alamos

National Laboratory (<http://hiv-web.lanl.gov/content/hiv-db/mainpage.html>)). The N27D mutant in our studies was quite defective in DNA binding and helix destabilization, but this is a much more drastic change than those changes noted above. Further, as mentioned earlier, the N27D mutant was not found to be defective in replication *in vivo* in H9 cells. Based on the very low helix destabilizing activity of N27D *in vitro*, it seems reasonable to hypothesize that the aggregation activity of HIV-1 NC is more critical than helix-destabilizing activity for viral viability. Preliminary results suggest that this may be the case for retroviruses in general as NC proteins from SIV show reduced destabilizing activity and MuLV NC shows very little (See Chapter 3). It is important to note here that the apparent ability of viruses containing the N27D mutation to replicate in cells does not mean that defects would not become apparent after several rounds of replication, or that the mutant is as fit as wt virus. Competition growth assays between the mutant and wt NC would have to be performed to assess relative fitness of the mutant.

Since NC is involved in several steps of the viral life cycle, a comprehensive analysis of how it works could enable us to potentially figure out new ways to target this protein. Designing molecules/compounds like nucleic acid analogues which would bind tightly to NC and prevent it from interacting with nucleic acids does seem like a promising avenue of research. A number of NC inhibitors have been proposed and some are undergoing initial phases of clinical trials. It is very probable that drug-resistant mutants would appear which are able to circumvent these drugs. Identifying amino acid residues that are invariant, would give us an idea of which ones are indispensable for viral viability. Among the two residues that were found to contribute significantly to HIV-1 NC's destabilizing activity, the asparagine residue at position 27 was found to be

fairly highly conserved and hence could be considered as generally invariant. If NC inhibitors could be designed that can specifically bind/disrupt this residue, perhaps the destabilizing activity of NC can be disturbed. In one report, threonine at position 24 of HIV-1 NC was changed to alanine; resulting in no loss in infectivity [140]. It should be noted that this NC protein had a threonine at position 24 rather than isoleucine as is present in the strain we used, but the fact that it could be changed to alanine without loss of infectivity suggests some flexibility at this position. This position is not as strongly conserved as the other four examined in our assays. Many HIV isolates have leucine at this position and some have threonine and valine. If these NC isolates could be examined *in vitro* and tested for their destabilizing and aggregation activities, we can possibly gain further insight into the apparent flexibility at this position. One possibility that is suggested by the viability in cells with the N27D mutation is that NC needs relatively little helix destabilizing activity to complete replication. Since many of the naturally occurring HIV isolates that have been sequenced have residues other than isoleucine at position 24, it would be interesting to test these isolates for helix destabilizing activity. If lower (or higher) activity is found this would suggest that the virus can survive with a variable level of helix destabilizing activity and this activity may not be a major determinant of viral fitness. With regard to position 24, we also test a glutamic acid mutant at this position (I24E, data not shown). This mutant was highly defective, even more so than the F16WI24QN27D triple mutant. The severe change of a hydrophobic to a negatively charged polar amino acid essentially produced a dead protein illustrating that this position is not completely flexible.

Despite over twenty years of intensive laboratory and clinical research, AIDS continues to be one of the world's deadliest scourges. In the US alone, AIDS has claimed more than 400,000 human lives and continues to be the third leading cause for mortality between the ages of 25-44, resulting in nearly 10,000 deaths every year. Since 1981, more than 25 million people worldwide have died of AIDS. Statistics reveal that at the end of 2005, a staggering 39 million people worldwide were living with HIV/AIDS. Globally, AIDS is the leading cause of deaths among infectious diseases and is the fourth leading cause of worldwide mortality. Every year nearly 2.6 million people worldwide die of AIDS [6]. Intensive research has had a great deal of benefits and has caused US AIDS related deaths to drop by nearly 80% since its peak at 1991, but of late; results from recent drug-therapy/vaccine trials are not so promising [157]. Recently, mutant strains resistant to triple drug therapy have been reported. The main challenges towards designing effective vaccines/drugs are the high recombinogenic potential and mutation rates of the virus. With no effective cure in sight, an exhaustive understanding of specific interactions occurring during viral replication could provide worthwhile information which can be used in the design of new means of fighting HIV. Work done in my thesis has given new insight towards understanding intricate mechanisms behind HIV replication and recombination. Molecular mechanisms behind interaction of HIV-1 NC with nucleic acids are also investigated in great detail. It also enhances one's perspective of how HIV-1 NC facilitates replication and recombination at an amino acid level. As described earlier, some significant gaps have been covered and future directions are clearer.

BIBLIOGRAPHY

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Retroviral Virions and Genomes*, in *Retroviruses*. 1997, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. p. 27-70.
2. *How HIV Causes AIDS*. 2004. National Institute of Allergy and Infectious Diseases. URL: <http://www.niaid.nih.gov/factsheets/howhiv.htm>.
3. *Timeline, 1981-1988*. National Institutes of Health. URL: <http://aidshistory.nih.gov/timeline/index.html>.
4. Kanabus, A. and Allen, S., *The Origins of HIV & the first cases of AIDS*. AVERT. URL: <http://www.avert.org/origins.htm>.
5. Noble, R., *HIV types, groups and subtypes*. AVERT. URL: <http://www.avert.org/hivtypes.htm>.
6. *Worldwide HIV & AIDS Statistics*. AVERT. URL: <http://www.avert.org/worldstats.htm>.
7. *What is AIDS?* AVERT. URL: <http://www.avert.org/aids.htm>.
8. *Stages of HIV infection*. AVERT. URL: <http://www.avert.org/hivstages.htm>.
9. *Introduction to HIV/AIDS Treatment*. AVERT. URL: <http://www.avert.org/introtrt.htm>.
10. Noble, R., *AIDS Vaccines and Microbicides*. AVERT. URL: <http://www.avert.org/vaccines-microbicides.htm>.
11. *HIV*. Wikipedia. URL: <http://en.wikipedia.org/wiki/HIV>.
12. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Reverse Transcriptase and the Generation of Retroviral DNA*, in *Retroviruses*. 1997, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. p. 121-160.

13. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Integration*, in *Retroviruses*. 1997, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. p. 121-160.
14. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Synthesis and Processing of Viral RNA*, in *Retroviruses*. 1997, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. p. 121-160.
15. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Synthesis, Assembly, and Processing of Viral Proteins*, in *Retroviruses*. 1997, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. p. 121-160.
16. Le Grice, S. F., Naas, T., Wohlgensinger, B., and Schatz, O., *Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase*. EMBO J., 1991. **10**(12): p. 3905-3911.
17. Hostomsky, Z., Hostomska, Z., Fu, T.-B., and Taylor, J., *Reverse Transcriptase of Human Immunodeficiency Virus Type 1: Functionality of Subunits of the Heterodimer in DNA Synthesis*. J. Virol., 1992. **66**(5): p. 3179-3182.
18. Larder, B. A., Purifoy, D. J., Powell, K. L., and Darby, G., *Site-specific mutagenesis of AIDS virus reverse transcriptase*. Nature, 1987. **327**(6124): p. 716-717.
19. Mizrahi, V., *Analysis of the ribonuclease H activity of HIV-1 reverse transcriptase using RNA-DNA hybrid substrates derived from the gag region of HIV-1*. Biochemistry (Mosc). 1989. **28**(23): p. 9088-9094.
20. Repaske, R., Hartley, J. W., Kavlick, M. F., O'Neill, R. R., and Austin, J. B., *Inhibition of RNase H activity and viral replication by single mutations in the 3' region of Moloney murine leukemia virus reverse transcriptase*. J. Virol., 1989. **63**(3): p. 1460-1464.
21. Arthur, A., *Overview of RT Structure*, HIV Drug Resistance Program.
22. Boyer, J. C., Bebenek, K., and Kunkel, T. A., *Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates*. Proc. Natl. Acad. Sci. U. S. A., 1992. **89**(15): p. 6919-6923.

23. Yu, H. and Goodman, M. F., *Comparison of HIV-1 and Avian Myeloblastosis Virus Reverse Transcriptase Fidelity on RNA and DNA Templates*. J. Biol. Chem., 1992. **267**(15): p. 10888-10896.
24. Wain-Hobson, S., *Human immunodeficiency virus type 1 quasispecies in vivo and ex vivo*. Curr. Top. Microbiol. Immunol., 1992. **176**: p. 181-193.
25. O'Neil, P. K., Sun, G., Yu, H., Ron, Y., Dougherty, J. P., and Preston, B. D., *Mutational analysis of HIV-1 long terminal repeats to explore the relative contribution of reverse transcriptase and RNA polymerase II to viral mutagenesis*. J. Biol. Chem., 2002. **277**(41): p. 38053-38061.
26. Anderson, J. A., Bowman, E. H., and Hu, W. S., *Retroviral recombination rates do not increase linearly with marker distance and are limited by the size of the recombining subpopulation*. J. Virol., 1998. **72**(2): p. 1195-1202.
27. Yu, H., Jetzt, A. E., Ron, Y., Preston, B. D., and Dougherty, J. P., *The nature of human immunodeficiency virus type 1 strand transfers*. J. Biol. Chem., 1998. **273**(43): p. 28384-28391.
28. Jetzt, A. E., Yu, H., Klarmann, G. J., Ron, Y., Preston, B. D., and Dougherty, J. P., *High rate of recombination throughout the human immunodeficiency virus type 1 genome*. J. Virol., 2000. **74**(3): p. 1234-1240.
29. Zhuang, J., Jetzt, A. E., Sun, G., Yu, H., Klarmann, G., Ron, Y., Preston, B. D., and Dougherty, J. P., *Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots*. J. Virol., 2002. **76**(22): p. 11273-11282.
30. Levy, D. N., Aldrovandi, G. M., Kutsch, O., and Shaw, G. M., *Dynamics of HIV-1 recombination in its natural target cells*. Proc. Natl. Acad. Sci. U. S. A., 2004. **101**(12): p. 4204-4209.
31. Coffin, J. M., *Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses*. J. Gen. Virol., 1979. **42**(1): p. 1-26.
32. Hu, W. S. and Temin, H. M., *Effect of gamma radiation on retroviral recombination*. J. Virol., 1992. **66**(7): p. 4457-4463.

33. Xu, H. and Boeke, J. D., *High-frequency deletion between homologous sequences during retrotransposition of Ty elements in Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A., 1987. **84**(23): p. 8553-8557.
34. DeStefano, J. J., Bambara, R. A., and Fay, P. J., *The mechanism of human immunodeficiency virus reverse transcriptase- catalyzed strand transfer from internal regions of heteropolymeric RNA templates*. J. Biol. Chem., 1994. **269**(1): p. 161-168.
35. Junghans, R. P., Boone, L. R., and Skalka, A. M., *Retroviral DNA H structures: displacement-assimilation model of recombination*. Cell, 1982. **30**(1): p. 53-62.
36. Charneau, P., Alizon, M., and Clavel, F., *A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication*. J. Virol., 1992. **66**(5): p. 2814-2820.
37. Darlix, J. L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P., *First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses*. J. Mol. Biol., 1995. **254**(4): p. 523-537.
38. Henderson, L. E., Bowers, M. A., Sowder II, R. C., Serabyn, S. A., Johnson, D. G., Bess Jr., J. W., Arthur, L. O., Bryant, D. K., and Fenselau, C., *Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: post translational modifications*. J. Virol., 1992. **4**: p. 1856-1865.
39. Tritch, R. J., Cheng, Y. E., Yin, F. H., and Erickson-Viitanen, S., *Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 gag polyprotein*. J. Virol., 1991. **65**(2): p. 922-930.
40. Coffin, J. M., Hughes, S. H., and Varmus, H. E., *Retroviruses*. 1997, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
41. Rein, A., Henderson, L. E., and Levin, J. G., *Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication*. Trends Biochem. Sci., 1998. **23**(8): p. 297-301.
42. Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J. L., and Roques, B. P., *Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR*. EMBO J., 1992. **11**(8): p. 3059-3065.

43. Surovoy, A., Dannull, J., Moelling, K., and Jung, G., *Conformational and nucleic acid binding studies on the synthetic nucleocapsid protein of HIV-1*. J. Mol. Biol., 1993. **229**(1): p. 94-104.
44. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F., *Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element*. Science, 1998. **279**(5349): p. 384-388.
45. DeStefano, J. J., *Interaction of human immunodeficiency virus nucleocapsid protein with a structure mimicking a replication intermediate. Effects on stability, reverse transcriptase binding, and strand transfer*. J. Biol. Chem., 1996. **271**(27): p. 16350-16356.
46. Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Levin, J. G., *Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer*. J. Virol., 2000. **74**(19): p. 8980-8988.
47. Hargittai, M. R., Mangla, A. T., Gorelick, R. J., and Musier-Forsyth, K., *HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys,3) structural changes but are not critical for primer/template annealing*. J. Mol. Biol., 2001. **312**(5): p. 985-997.
48. Schmalzbauer, E., Strack, B., Dannull, J., Guehmann, S., and Moelling, K., *Mutations of basic amino acids of NCp7 of human immunodeficiency virus type 1 affect RNA binding in vitro*. J. Virol., 1996. **70**(2): p. 771-777.
49. Berthoux, L., Pechoux, C., and Darlix, J. L., *Multiple effects of an anti-human immunodeficiency virus nucleocapsid inhibitor on virus morphology and replication*. J. Virol., 1999. **73**(12): p. 10000-10009.
50. Rice, W. G., Supko, J. G., Malspeis, L., Buckheit, R. W., Jr., Clanton, D., Bu, M., Graham, L., Schaeffer, C. A., Turpin, J. A., Domagala, J., Gogliotti, R., Bader, J. P., Halliday, S. M., Coren, L., Sowder, R. C., 2nd, Arthur, L. O., and Henderson, L. E., *Inhibitors of HIV nucleocapsid protein zinc fingers as candidates for the treatment of AIDS*. Science, 1995. **270**(5239): p. 1194-1197.
51. Nagy, K., Young, M., Baboonian, C., Merson, J., Whittle, P., and Oroszlan, S., *Antiviral activity of human immunodeficiency virus type 1 protease inhibitors in a single cycle of infection: evidence for a role of protease in the early phase*. J. Virol., 1994. **68**(2): p. 757-765.

52. Rossio, J. L., Esser, M. T., Suryanarayana, K., Schneider, D. K., Bess, J. W., Jr., Vasquez, G. M., Wiltout, T. A., Chertova, E., Grimes, M. K., Sattentau, Q., Arthur, L. O., Henderson, L. E., and Lifson, J. D., *Inactivation of human immunodeficiency virus type 1 infectivity with preservation of conformational and functional integrity of virion surface proteins*. J. Virol., 1998. **72**(10): p. 7992-8001.
53. Lapadat-Tapolsky, M., De Rocquigny, H., Van Gent, D., Roques, B., Plasterk, R., and Darlix, J. L., *Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle [published erratum appears in Nucleic Acids Res 1993 Apr 25;21(8):2024]*. Nucleic Acids Res, 1993. **21**(4): p. 831-839.
54. Karpel, R. L., Henderson, L. E., and Oroszlan, S., *Interactions of retroviral structural proteins with single-stranded nucleic acids*. J. Biol. Chem., 1987. **262**(11): p. 4961-4967.
55. Khan, R. and Giedroc, D. P., *Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA*. J. Biol. Chem., 1992. **267**(10): p. 6689-6695.
56. Remy, E., de Rocquigny, H., Petitjean, P., Muriaux, D., Theilleux, V., Paoletti, J., and Roques, B. P., *The annealing of tRNA^{3Lys} to human immunodeficiency virus type 1 primer binding site is critically dependent on the NCp7 zinc fingers structure*. J. Biol. Chem., 1998. **273**(9): p. 4819-4822.
57. De Rocquigny, H., Gabus, C., Vincent, A., Fournie-Zaluski, M. C., Roques, B., and Darlix, J. L., *Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers*. Proc. Natl. Acad. Sci. U. S. A., 1992. **89**(14): p. 6472-6476.
58. Barat, C., Lullien, V., Schatz, O., Keith, G., Nugeyre, M. T., Gruninger-Leitch, F., Barre-Sinoussi, F., LeGrice, S. F., and Darlix, J. L., *HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA*. EMBO J., 1989. **8**(11): p. 3279-3285.
59. De Rocquigny, H., Ficheux, D., Gabus, C., Allain, B., Fournie-Zaluski, M. C., Darlix, J. L., and Roques, B. P., *Two short basic sequences surrounding the zinc finger of nucleocapsid protein NCp10 of Moloney murine leukemia virus are critical for RNA annealing activity*. Nucleic Acids Res, 1993. **21**(4): p. 823-829.

60. Prats, A. C., Housset, V., de Billy, G., Cornille, F., Prats, H., Roques, B., and Darlix, J. L., *Viral RNA annealing activities of the nucleocapsid protein of Moloney murine leukemia virus are zinc independent*. Nucleic Acids Res, 1991. **19**(13): p. 3533-3541.
61. Ji, X., Klarmann, G. J., and Preston, B. D., *Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro*. Biochemistry (Mosc). 1996. **35**: p. 132-143.
62. Lener, D., Tanchou, V., Roques, B. P., Le Grice, S. F., and Darlix, J. L., *Involvement of HIV-1 nucleocapsid protein in the recruitment of reverse transcriptase into nucleoprotein complexes formed in vitro*. J. Biol. Chem., 1998. **273**(50): p. 33781-33786.
63. Rodriguez-Rodriguez, L., Tsuchihashi, Z., Fuentes, G. M., Bambara, R. A., and Fay, P. J., *Influence of human immunodeficiency virus nucleocapsid protein on synthesis and strand transfer by the reverse transcriptase in vitro*. J. Biol. Chem., 1995. **270**(25): p. 15005-15011.
64. Chen, Y., Balakrishnan, M., Roques, B. P., and Bambara, R. A., *Steps of the acceptor invasion mechanism for HIV-1 minus strand strong stop transfer*. J. Biol. Chem., 2003. **278**(40): p. 38368-38375.
65. Wisniewski, M., Chen, Y., Balakrishnan, M., Palaniappan, C., Roques, B. P., Fay, P. J., and Bambara, R. A., *Substrate requirements for secondary cleavage by HIV-1 reverse transcriptase RNase H*. J. Biol. Chem., 2002. **277**(32): p. 28400-28410.
66. Allain, B., Lapadat-Tapolsky, M., Berlioz, C., and Darlix, J. L., *Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome*. EMBO J., 1994. **13**(4): p. 973-981.
67. DeStefano, J. J., *Human immunodeficiency virus nucleocapsid protein stimulates strand transfer from internal regions of heteropolymeric RNA templates*. Arch. Virol., 1995. **140**(10): p. 1775-1789.
68. Guo, J., Henderson, L. E., Bess, J., Kane, B., and Levin, J. G., *Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA*. J. Virol., 1997. **71**(7): p. 5178-5188.

69. Kim, J. K., Palaniappan, C., Wu, W., Fay, P. J., and Bambara, R. A., *Evidence for a unique mechanism of strand transfer from the transactivation response region of HIV-1*. J. Biol. Chem., 1997. **272**(27): p. 16769-16777.
70. Peliska, J. A., Balasubramanian, S., Giedroc, D. P., and Benkovic, S. J., *Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity*. Biochemistry (Mosc). 1994. **33**(46): p. 13817-13823.
71. Raja, A. and DeStefano, J. J., *Kinetic analysis of the effect of HIV nucleocapsid protein (NCp) on internal strand transfer reactions*. Biochemistry (Mosc). 1999. **38**(16): p. 5178-5184.
72. You, J. C. and McHenry, C. S., *Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription*. J. Biol. Chem., 1994. **269**(50): p. 31491-31495.
73. Wu, T., Guo, J., Bess, J., Henderson, L. E., and Levin, J. G., *Molecular requirements for human immunodeficiency virus type 1 plus-strand transfer: analysis in reconstituted and endogenous reverse transcription systems*. J. Virol., 1999. **73**(6): p. 4794-4805.
74. Carteau, S., Batson, S. C., Poljak, L., Mouscadet, J. F., de Rocquigny, H., Darlix, J. L., Roques, B. P., Kas, E., and Auclair, C., *Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg²⁺-dependent DNA integration in vitro*. J. Virol., 1997. **71**(8): p. 6225-6229.
75. Carteau, S., Gorelick, R. J., and Bushman, F. D., *Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro: stimulation by the viral nucleocapsid protein*. J. Virol., 1999. **73**(8): p. 6670-6679.
76. Buckman, J. S., Bosche, W. J., and Gorelick, R. J., *Human immunodeficiency virus type 1 nucleocapsid zn(2+) fingers are required for efficient reverse transcription, initial integration processes, and protection of newly synthesized viral DNA*. J. Virol., 2003. **77**(2): p. 1469-1480.
77. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., Chancellor, K. J., Wu, Z. R., and Summers, M. F., *NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition*. J. Mol. Biol., 2000. **301**(2): p. 491-511.

78. Gorelick, R. J., Gagliardi, T. D., Bosche, W. J., Wilttrout, T. A., Coren, L. V., Chabot, D. J., Lifson, J. D., Henderson, L. E., and Arthur, L. O., *Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences*. Virology, 1999. **256**(1): p. 92-104.
79. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A., *Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a "zinc finger-like" protein sequence*. Proc. Natl. Acad. Sci. U. S. A., 1988. **85**(22): p. 8420-8424.
80. Poon, D. T., Wu, J., and Aldovini, A., *Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity*. J. Virol., 1996. **70**(10): p. 6607-6616.
81. Poon, D. T., Li, G., and Aldovini, A., *Nucleocapsid and matrix protein contributions to selective human immunodeficiency virus type 1 genomic RNA packaging*. J. Virol., 1998. **72**(3): p. 1983-1993.
82. Levin, J. G., Guo, J., Rouzina, I., and Musier-Forsyth, K., *Nucleic Acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism*. Prog. Nucleic Acids Res. Mol. Biol., 2005. **80**: p. 217-286.
83. Feng, Y. X., Copeland, T. D., Henderson, L. E., Gorelick, R. J., Bosche, W. J., Levin, J. G., and Rein, A., *HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro*. Proc. Natl. Acad. Sci. U. S. A., 1996. **93**(15): p. 7577-7581.
84. Fu, W., Gorelick, R. J., and Rein, A., *Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions*. J. Virol., 1994. **68**(8): p. 5013-5018.
85. Fu, W. and Rein, A., *Maturation of dimeric viral RNA of Moloney murine leukemia virus*. J. Virol., 1993. **67**(9): p. 5443-5449.
86. Herschlag, D., *RNA chaperones and the RNA folding problem*. J. Biol. Chem., 1995. **270**(36): p. 20871-20874.

87. Golinelli, M. P. and Hughes, S. H., *Secondary structure in the nucleic acid affects the rate of HIV-1 nucleocapsid-mediated strand annealing*. Biochemistry (Mosc). 2003. **42**(27): p. 8153-8162.
88. Heath, M. J. and Destefano, J. J., *A complementary single-stranded docking site is required for enhancement of strand exchange by human immunodeficiency virus nucleocapsid protein on substrates that model viral recombination*. Biochemistry (Mosc). 2005. **44**(10): p. 3915-3925.
89. Tsuchihashi, Z. and Brown, P. O., *DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein*. J. Virol., 1994. **68**(9): p. 5863-5870.
90. Dib-Hajj, F., Khan, R., and Giedroc, D. P., *Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity*. Protein Sci., 1993. **2**(2): p. 231-243.
91. Heath, M. J., Derebail, S. S., Gorelick, R. J., and DeStefano, J. J., *Differing roles of the N-terminal and C-terminal zinc fingers in HIV-1 nucleocapsid protein enhanced nucleic acid annealing*. J. Biol. Chem., 2003. **278**: p. 30755-30763.
92. Christiansen, C. and Baldwin, R. L., *Catalysis of DNA reassociation by the Escherichia coli DNA binding protein: A polyamine-dependent reaction*. J. Mol. Biol., 1977. **115**(3): p. 441-454.
93. Pontius, B. W. and Berg, P., *Rapid assembly and disassembly of complementary DNA strands through an equilibrium intermediate state mediated by A1 hnRNP protein*. J. Biol. Chem., 1992. **267**(20): p. 13815-13818.
94. Pant, K., Karpel, R. L., and Williams, M. C., *Kinetic regulation of single DNA molecule denaturation by T4 gene 32 protein structural domains*. J. Mol. Biol., 2003. **327**(3): p. 571-578.
95. Hargittai, M. R., Gorelick, R. J., Rouzina, I., and Musier-Forsyth, K., *Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site*. J. Mol. Biol., 2004. **337**(4): p. 951-968.
96. South, T. L., Blake, P. R., Hare, D. R., and Summers, M. F., *C-terminal retroviral-type zinc finger domain from the HIV-1 nucleocapsid protein is*

- structurally similar to the N-terminal zinc finger domain.* Biochemistry (Mosc). 1991. **30**(25): p. 6342-6349.
97. South, T. L., Blake, P. R., Sowder, R. C., 3rd, Arthur, L. O., Henderson, L. E., and Summers, M. F., *The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers.* Biochemistry (Mosc). 1990. **29**(34): p. 7786-7789.
 98. Guo, J., Wu, T., Kane, B. F., Johnson, D. G., Henderson, L. E., Gorelick, R. J., and Levin, J. G., *Subtle alterations of the native zinc finger structures have dramatic effects on the nucleic acid chaperone activity of human immunodeficiency virus type 1 nucleocapsid protein.* J. Virol., 2002. **76**(9): p. 4370-4378.
 99. Gorelick, R. J., Nigida, S. M., Jr., Bess, J. W., Jr., Arthur, L. O., Henderson, L. E., and Rein, A., *Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA.* J. Virol., 1990. **64**(7): p. 3207-3211.
 100. Williams, M. C., Rouzina, I., Wenner, J. R., Gorelick, R. J., Musier-Forsyth, K., and Bloomfield, V. A., *Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching.* Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(11): p. 6121-6126.
 101. Urbaneja, M. A., Wu, M., Casas-Finet, J. R., and Karpel, R. L., *HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity.* J. Mol. Biol., 2002. **318**(3): p. 749-764.
 102. Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E., *Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides.* J. Virol., 1998. **72**(3): p. 1902-1909.
 103. Berglund, J. A., Charpentier, B., and Rosbash, M., *A high affinity binding site for the HIV-1 nucleocapsid protein.* Nucleic Acids Res, 1997. **25**(5): p. 1042-1049.
 104. D'Souza, V. and Summers, M. F., *Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus.* Nature, 2004. **431**(7008): p. 586-590.

105. Mely, Y., de Rocquigny, H., Sorinas-Jimeno, M., Keith, G., Roques, B. P., Marquet, R., and Gerard, D., *Binding of the HIV-1 nucleocapsid protein to the primer tRNA(3Lys), in vitro, is essentially not specific*. J. Biol. Chem., 1995. **270**(4): p. 1650-1656.
106. Khan, R. and Giedroc, D. P., *Nucleic acid binding properties of recombinant Zn2 HIV-1 nucleocapsid protein are modulated by COOH-terminal processing*. J. Biol. Chem., 1994. **269**(36): p. 22538-22546.
107. You, J. C. and McHenry, C. S., *HIV nucleocapsid protein. Expression in Escherichia coli, purification, and characterization*. J. Biol. Chem., 1993. **268**(22): p. 16519-16527.
108. Urbaneja, M. A., Kane, B. P., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Casas-Finet, J. R., *Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: elucidation of the structural determinants for function*. J. Mol. Biol., 1999. **287**(1): p. 59-75.
109. Johnson, P. E., Turner, R. B., Wu, Z. R., Hairston, L., Guo, J., Levin, J. G., and Summers, M. F., *A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription*. Biochemistry (Mosc). 2000. **39**(31): p. 9084-9091.
110. Tisne, C., Roques, B. P., and Dardel, F., *Heteronuclear NMR studies of the interaction of tRNA(Lys)3 with HIV-1 nucleocapsid protein*. J. Mol. Biol., 2001. **306**(3): p. 443-454.
111. Tisne, C., Roques, B. P., and Dardel, F., *Specific recognition of primer tRNA Lys 3 by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension*. Biochimie, 2003. **85**(5): p. 557-561.
112. Tisne, C., Roques, B. P., and Dardel, F., *The annealing mechanism of HIV-1 reverse transcription primer onto the viral genome*. J. Biol. Chem., 2004. **279**(5): p. 3588-3595.
113. Ramboarina, S., Srividya, N., Atkinson, R. A., Morellet, N., Roques, B. P., Lefevre, J. F., Mely, Y., and Kieffer, B., *Effects of temperature on the dynamic behaviour of the HIV-1 nucleocapsid NCp7 and its DNA complex*. J. Mol. Biol., 2002. **316**(3): p. 611-627.

114. Stote, R. H., Kellenberger, E., Muller, H., Bombarda, E., Roques, B. P., Kieffer, B., and Mely, Y., *Structure of the His44 --> Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein: a combined NMR, molecular dynamics simulation, and fluorescence study*. Biochemistry (Mosc). 2004. **43**(24): p. 7687-7697.
115. Vuilleumier, C., Bombarda, E., Morellet, N., Gerard, D., Roques, B. P., and Mely, Y., *Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study*. Biochemistry (Mosc). 1999. **38**(51): p. 16816-16825.
116. Shubsda, M. F., Paoletti, A. C., Hudson, B. S., and Borer, P. N., *Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein*. Biochemistry (Mosc). 2002. **41**(16): p. 5276-5282.
117. Cantor, C. R. and Schimmel, P. R., *Biophysical Chemistry. Part III. The Behavior of Biological Macromolecules*. 1980, San Francisco, CA: W. H. Freeman & Co.
118. Bloomfield, V. A., Crothers, D. M., and Tinoco, J. I., *Nucleic Acids: Structures, Properties and Functions*. 1998, Mill Valley, CA: University Science Press.
119. Bloomfield, V. A., *DNA condensation*. Curr. Opin. Struct. Biol., 1996. **6**(3): p. 334-341.
120. Paoletti, A. C., Shubsda, M. F., Hudson, B. S., and Borer, P. N., *Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1*. Biochemistry (Mosc). 2002. **41**(51): p. 15423-15428.
121. Maki, A. H., Ozarowski, A., Misra, A., Urbaneja, M. A., and Casas-Finet, J. R., *Phosphorescence and optically detected magnetic resonance of HIV-1 nucleocapsid protein complexes with stem-loop sequences of the genomic Psi-recognition element*. Biochemistry (Mosc). 2001. **40**(5): p. 1403-1412.
122. Hagan, N. and Fabris, D., *Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element*. Biochemistry (Mosc). 2003. **42**(36): p. 10736-10745.
123. Khandogin, J., Musier-Forsyth, K., and York, D. M., *Insights into the regioselectivity and RNA-binding affinity of HIV-1 nucleocapsid protein from linear-scaling quantum methods*. J. Mol. Biol., 2003. **330**(5): p. 993-1004.

124. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., and Summers, M. F., *NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform*. J. Mol. Biol., 2000. **299**(1): p. 145-156.
125. Amarasinghe, G. K., Zhou, J., Miskimon, M., Chancellor, K. J., McDonald, J. A., Matthews, A. G., Miller, R. R., Rouse, M. D., and Summers, M. F., *Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein. structural studies and implications for genome recognition*. J. Mol. Biol., 2001. **314**(5): p. 961-970.
126. Clever, J., Sasseti, C., and Parslow, T. G., *RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1*. J. Virol., 1995. **69**(4): p. 2101-2109.
127. Panganiban, A. T. and Fiore, D., *Ordered interstrand and intrastrand DNA transfer during reverse transcription*. Science, 1988. **241**(4869): p. 1064-1069.
128. Hu, W. S. and Temin, H. M., *Retroviral recombination and reverse transcription*. Science, 1990. **250**(4985): p. 1227-1233.
129. van Wamel, J. L. and Berkhout, B., *The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly*. Virology, 1998. **244**(2): p. 245-251.
130. Williams, M. C., Gorelick, R. J., and Musier-Forsyth, K., *Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function*. Proc. Natl. Acad. Sci. U. S. A., 2002. **99**(13): p. 8614-8619.
131. Lapadat-Tapolsky, M., Gabus, C., Rau, M., and Darlix, J. L., *Possible roles of HIV-1 nucleocapsid protein in the specificity of proviral DNA synthesis and in its variability*. J. Mol. Biol., 1997. **268**(2): p. 250-260.
132. Driscoll, M. D., Golinelli, M. P., and Hughes, S. H., *In vitro analysis of human immunodeficiency virus type 1 minus-strand strong-stop DNA synthesis and genomic RNA processing*. J. Virol., 2001. **75**(2): p. 672-686.
133. Heilman-Miller, S. L., Wu, T., and Levin, J. G., *Alteration of Nucleic Acid Structure and Stability Modulates the Efficiency of Minus-Strand Transfer Mediated by the HIV-1 Nucleocapsid Protein*. J. Biol. Chem., 2004. **279**(42): p. 44154-44165.

134. Hong, M. K., Harbron, E. J., O'Connor, D. B., Guo, J., Barbara, P. F., Levin, J. G., and Musier-Forsyth, K., *Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer*. J. Mol. Biol., 2003. **325**(1): p. 1-10.
135. Derebail, S. S., Heath, M. J., and DeStefano, J. J., *Evidence for the differential effects of nucleocapsid protein on strand transfer in various regions of the HIV genome*. J. Biol. Chem., 2003. **278**: p. 15702-15712.
136. Wu, W., Henderson, L. E., Copeland, T. D., Gorelick, R. J., Bosche, W. J., Rein, A., and Levin, J. G., *Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract*. J. Virol., 1996. **70**(10): p. 7132-7142.
137. Derebail, S. S. and Destefano, J. J., *Mechanistic Analysis of Pause Site-dependent and -independent Recombinogenic Strand Transfer from Structurally Diverse Regions of the HIV Genome*. J. Biol. Chem., 2004. **279**(46): p. 47446-47454.
138. Beltz, H., Clauss, C., Piemont, E., Ficheux, D., Gorelick, R. J., Roques, B., Gabus, C., Darlix, J. L., de Rocquigny, H., and Mely, Y., *Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis*. J. Mol. Biol., 2005. **348**(5): p. 1113-1126.
139. McGrath, C. F., Buckman, J. S., Gagliardi, T. D., Bosche, W. J., Coren, L. V., and Gorelick, R. J., *Human cellular nucleic acid-binding protein Zn²⁺ fingers support replication of human immunodeficiency virus type 1 when they are substituted in the nucleocapsid protein*. J. Virol., 2003. **77**(15): p. 8524-8531.
140. Dorfman, T., Luban, J., Goff, S. P., Haseltine, W. A., and Gottlinger, H. G., *Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein*. J. Virol., 1993. **67**(10): p. 6159-6169.
141. Mark-Danieli, M., Laham, N., Kenan-Eichler, M., Castiel, A., Melamed, D., Landau, M., Bouvier, N. M., Evans, M. J., and Bacharach, E., *Single point mutations in the zinc finger motifs of the human immunodeficiency virus type 1 nucleocapsid alter RNA binding specificities of the gag protein and enhance packaging and infectivity*. J. Virol., 2005. **79**(12): p. 7756-7767.

142. Yovandich, J. L., Chertova, E. N., Kane, B. P., Gagliardi, T. D., Bess, J. W., Jr., Sowder, R. C., 2nd, Henderson, L. E., and Gorelick, R. J., *Alteration of zinc-binding residues of simian immunodeficiency virus p8(NC) results in subtle differences in gag processing and virion maturation associated with degradative loss of mutant NC*. J. Virol., 2001. **75**(1): p. 115-124.
143. Gorelick, R. J., Benveniste, R. E., Lifson, J. D., Yovandich, J. L., Morton, W. R., Kuller, L., Flynn, B. M., Fisher, B. A., Rossio, J. L., Piatak, M., Jr., Bess, J. W., Jr., Henderson, L. E., and Arthur, L. O., *Protection of Macaca nemestrina from disease following pathogenic simian immunodeficiency virus (SIV) challenge: utilization of SIV nucleocapsid mutant DNA vaccines with and without an SIV protein boost*. J. Virol., 2000. **74**(24): p. 11935-11949.
144. Gorelick, R. J., Lifson, J. D., Yovandich, J. L., Rossio, J. L., Piatak, M., Jr., Scarzello, A. J., Knott, W. B., Bess, J. W., Jr., Fisher, B. A., Flynn, B. M., Henderson, L. E., Arthur, L. O., and Benveniste, R. E., *Mucosal challenge of Macaca nemestrina with simian immunodeficiency virus (SIV) following SIV nucleocapsid mutant DNA vaccination*. J. Med. Primatol., 2000. **29**(3-4): p. 209-219.
145. Gorelick, R. J., Benveniste, R. E., Gagliardi, T. D., Wilttrout, T. A., Busch, L. K., Bosche, W. J., Coren, L. V., Lifson, J. D., Bradley, P. J., Henderson, L. E., and Arthur, L. O., *Nucleocapsid protein zinc-finger mutants of simian immunodeficiency virus strain mne produce virions that are replication defective in vitro and in vivo*. Virology, 1999. **253**(2): p. 259-270.
146. Urbaneja, M. A., McGrath, C. F., Kane, B. P., Henderson, L. E., and Casas-Finet, J. R., *Nucleic acid binding properties of the simian immunodeficiency virus nucleocapsid protein NCp8*. J. Biol. Chem., 2000. **275**(14): p. 10394-10404.
147. Morellet, N., Meudal, H., Bouaziz, S., and Roques, B. P., *Structure of the zinc finger domain encompassing residues 13-51 of the nucleocapsid protein from simian immunodeficiency virus*. Biochem. J., 2006. **393**(Pt 3): p. 725-732.
148. Zhang, W. H., Hwang, C. K., Hu, W. S., Gorelick, R. J., and Pathak, V. K., *Zinc finger domain of murine leukemia virus nucleocapsid protein enhances the rate of viral DNA synthesis in vivo*. J. Virol., 2002. **76**(15): p. 7473-7484.
149. Gorelick, R. J., Fu, W., Gagliardi, T. D., Bosche, W. J., Rein, A., Henderson, L. E., and Arthur, L. O., *Characterization of the block in replication of nucleocapsid*

- protein zinc finger mutants from Moloney murine leukemia virus.* J. Virol., 1999. **73**(10): p. 8185-8195.
150. Gorelick, R. J., Chabot, D. J., Ott, D. E., Gagliardi, T. D., Rein, A., Henderson, L. E., and Arthur, L. O., *Genetic analysis of the zinc finger in the Moloney murine leukemia virus nucleocapsid domain: replacement of zinc-coordinating residues with other zinc-coordinating residues yields noninfectious particles containing genomic RNA.* J. Virol., 1996. **70**(4): p. 2593-2597.
 151. Rein, A., Harvin, D. P., Mirro, J., Ernst, S. M., and Gorelick, R. J., *Evidence that a central domain of nucleocapsid protein is required for RNA packaging in murine leukemia virus.* J. Virol., 1994. **68**(9): p. 6124-6129.
 152. D'Souza, V., Melamed, J., Habib, D., Pullen, K., Wallace, K., and Summers, M. F., *Identification of a high affinity nucleocapsid protein binding element within the Moloney murine leukemia virus Psi-RNA packaging signal: implications for genome recognition.* J. Mol. Biol., 2001. **314**(2): p. 217-232.
 153. Allain, B., Rascle, J. B., de Rocquigny, H., Roques, B., and Darlix, J. L., *CIS elements and trans-acting factors required for minus strand DNA transfer during reverse transcription of the genomic RNA of murine leukemia virus.* J. Mol. Biol., 1998. **277**(2): p. 225-235.
 154. Rein, A., Ott, D. E., Mirro, J., Arthur, L. O., Rice, W., and Henderson, L. E., *Inactivation of murine leukemia virus by compounds that react with the zinc finger in the viral nucleocapsid protein.* J. Virol., 1996. **70**(8): p. 4966-4972.
 155. Turpin, J. A., Terpening, S. J., Schaeffer, C. A., Yu, G., Glover, C. J., Felsted, R. L., Sausville, E. A., and Rice, W. G., *Inhibitors of human immunodeficiency virus type 1 zinc fingers prevent normal processing of gag precursors and result in the release of noninfectious virus particles.* J. Virol., 1996. **70**(9): p. 6180-6189.
 156. Goldschmidt, V., Didierjean, J., Ehresmann, B., Ehresmann, C., Isel, C., and Marquet, R., *Mg²⁺ dependency of HIV-1 reverse transcription, inhibition by nucleoside analogues and resistance.* Nucleic Acids Res., 2006. **34**(1): p. 42-52.
 157. Noble, R., *United States Statistics Summary.* AVERT. URL: <http://www.avert.org/statsum.htm>.